

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 24, 2006, 05:46:37 ; Search time 851.046 Seconds
(without alignments)
2103.915 Million cell updates/sec

Title: US-10-601-913-1

Perfect score: 28

Sequence: 1 GACATTATGTTATAGTTGATGAGAC 28

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 6366136 seqs, 31973710525 residues

Total number of hits satisfying chosen parameters: 12732272

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 1000 summaries

Database :

GenEmbl:*
1: gb env:*
2: gb pac:*
3: gb_ph:*
4: gb_pl:*
5: gb_pr:*
6: gb_ro:*
7: gb_strs:*
8: gb_sy:*
9: gb_un:*
10: gb_vl:*
11: gb_ov:*
12: gb_hcg:*
13: gb_in:*
14: gb_cm:*
15: gb_ba:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	28	100.0	28	2	AR349061 Sequence
2	28	100.0	28	2	AR349062 Sequence
3	28	100.0	28	2	AR349063 Sequence
4	28	100.0	28	2	AR349064 Sequence
5	28	100.0	87	2	AR577462 Sequence
6	28	100.0	87	2	AX057044 Sequence
7	28	100.0	271	10	HPU14512 Sequence
8	28	100.0	272	10	HPU14513 Sequence
9	28	100.0	370	10	AF404692 Sequence
10	28	100.0	421	10	AF404695 Sequence
11	28	100.0	451	10	AF404703 Sequence
12	28	100.0	451	10	HPU34111 Sequence
13	28	100.0	456	10	HPU34115 Sequence
14	28	100.0	456	10	HPU34120 Sequence
15	28	100.0	456	10	HPU34121 Sequence
16	28	100.0	456	10	HPU34123 Sequence
17	28	100.0	456	10	HPU34124 Sequence
18	28	100.0	456	10	HPU34124 Sequence

19	28	100.0	456	10	HPU34126	U34126 Human papil
20	28	100.0	456	10	HPU34129	U34129 Human papil
21	28	100.0	458	10	AF003016	AF003016 Human pap
22	28	100.0	477	2	AR577459	AR577459 Sequence
23	28	100.0	477	2	AX057038	AX057038 Sequence
24	28	100.0	477	2	AX766063	AX766063 Sequence
25	28	100.0	477	10	AF486299	AF486299 Human pap
26	28	100.0	477	10	AF486300	AF486300 Human pap
27	28	100.0	477	10	AF486301	AF486301 Human pap
28	28	100.0	477	10	AF486302	AF486302 Human pap
29	28	100.0	477	10	AF486303	AF486303 Human pap
30	28	100.0	477	10	AF486304	AF486304 Human pap
31	28	100.0	477	10	AF486305	AF486305 Human pap
32	28	100.0	477	10	AF486306	AF486306 Human pap
33	28	100.0	477	10	AF486307	AF486307 Human pap
34	28	100.0	477	10	AF486308	AF486308 Human pap
35	28	100.0	477	10	AF486309	AF486309 Human pap
36	28	100.0	477	10	AF486310	AF486310 Human pap
37	28	100.0	477	10	AF486311	AF486311 Human pap
38	28	100.0	477	10	AF486312	AF486312 Human pap
39	28	100.0	477	10	AF486313	AF486313 Human pap
40	28	100.0	477	10	AF486315	AF486315 Human pap
41	28	100.0	477	10	AF486316	AF486316 Human pap
42	28	100.0	477	10	AF486317	AF486317 Human pap
43	28	100.0	477	10	AF486318	AF486318 Human pap
44	28	100.0	477	10	AF486319	AF486319 Human pap
45	28	100.0	477	10	AF486320	AF486320 Human pap
46	28	100.0	477	10	AF486321	AF486321 Human pap
47	28	100.0	477	10	AF486323	AF486323 Human pap
48	28	100.0	488	10	AF404702	AF404702 Human pap
49	28	100.0	519	2	AR095319	AR095319 Sequence
50	28	100.0	519	2	AR095320	AR095320 Sequence
51	28	100.0	519	2	AR173472	AR173472 Sequence
52	28	100.0	519	2	AR173473	AR173473 Sequence
53	28	100.0	540	10	AF404705	AF404705 Human pap
54	28	100.0	543	10	AF404693	AF404693 Human pap
55	28	100.0	543	10	AF404697	AF404697 Human pap
56	28	100.0	556	10	AY098919	AY098919 Human pap
57	28	100.0	556	10	AF548023	AF548023 Human pap
58	28	100.0	654	10	AF469198	AF469198 Human pap
59	28	100.0	776	2	119491	Sequence 1
60	28	100.0	776	2	AR199235	AR199235 Sequence
61	28	100.0	780	10	AF187865	AF187865 Human pap
62	28	100.0	780	10	AF187867	AF187867 Human pap
63	28	100.0	780	10	AF187868	AF187868 Human pap
64	28	100.0	780	10	AF187869	AF187869 Human pap
65	28	100.0	790	2	188813	Sequence 7
66	28	100.0	801	2	AR095318	AR095318 Sequence
67	28	100.0	801	2	AR173471	AR173471 Sequence
68	28	100.0	822	2	A98742	A98742 Sequence 3
69	28	100.0	822	2	BD080312	BD080312 Vaccine.
70	28	100.0	822	2	BD103161	BD103161 Vaccine.
71	28	100.0	822	2	AR183573	AR183573 Sequence
72	28	100.0	822	2	AX020926	AX020926 Sequence
73	28	100.0	879	2	A98748	A98748 Sequence 9
74	28	100.0	879	2	BD080315	BD080315 Vaccine.
75	28	100.0	879	2	BD103164	BD103164 Vaccine.
76	28	100.0	879	2	AR183576	AR183576 Sequence
77	28	100.0	879	2	AX020932	AX020932 Sequence
78	28	100.0	939	2	AX403959	AX403959 Sequence
79	28	100.0	939	2	AX460903	AX460903 Sequence
80	28	100.0	1000	2	A07622	A07622 Amplified s
81	28	100.0	1000	2	118920	Sequence 3
82	28	100.0	1116	2	A98744	A98744 Sequence 5
83	28	100.0	1116	2	BD080313	BD080313 Vaccine.
84	28	100.0	1116	2	BD103162	BD103162 Vaccine.
85	28	100.0	1116	2	AR183574	AR183574 Sequence
86	28	100.0	1116	2	AX020928	AX020928 Sequence
87	28	100.0	1173	2	A98752	A98752 Sequence 13
88	28	100.0	1173	2	BD080317	BD080317 Vaccine.
89	28	100.0	1173	2	BD103166	BD103166 Vaccine.
90	28	100.0	1173	2	AR183578	AR183578 Sequence
91	28	100.0	1173	2	AX020936	AX020936 Sequence

92	28	100.0	1267	10	HPA388056	AJ388056 Human pap	165	26.4	94.3	3531	10	AF001599	AF001599 Human pap
93	28	100.0	1267	10	HPA388058	AJ388058 Human pap	166	26.4	94.3	7904	2	I70143	I70143 Sequence 2
94	28	100.0	1267	10	HPA388059	AJ388059 Human pap	167	26.4	94.3	7904	10	AF125673	AF125673 Human pap
95	28	100.0	1267	10	HPA388060	AJ388060 Human pap	168	26.4	94.3	7904	10	AF472509	AF472509 Human pap
96	28	100.0	1267	10	HPA388061	AJ388061 Human pap	169	26.4	94.3	7904	10	AF536179	AF536179 Human pap
97	28	100.0	1267	10	HPA388066	AJ388066 Human pap	170	26.4	94.3	7905	10	HPU89348	HPU89348 Human pap
98	28	100.0	7902	2	AX800450	AX800450 Sequence	171	26.4	94.3	7906	10	AF536180	AF536180 Human pap
99	28	100.0	7904	2	BD070940	BD070940 Tissue sp	172	26.4	94.3	7906	10	AY66580	AY66580 Human pap
100	28	100.0	7904	2	BD225188	BD225188 Human pap	173	26.4	94.3	7907	10	AY66583	AY66583 Human pap
101	28	100.0	7904	2	CS073238	CS073238 Sequence	174	26.4	94.3	7908	10	AF472508	AF472508 Human pap
102	28	100.0	7904	2	CS195904	CS195904 Sequence	175	24.8	88.6	245	10	HPU14515	HPU14515
103	28	100.0	7904	2	I70142	I70142 Sequence 1	176	24.8	88.6	274	10	HPU14516	HPU14516
104	28	100.0	7904	2	I72345	I72345 Sequence 1	177	24.8	88.6	456	2	AR167393	AR167393 Sequence
105	28	100.0	7904	2	AX763489	AX763489 Sequence	178	24.8	88.6	456	10	HPU34107	HPU34107
106	28	100.0	7904	10	PPH16	KO2718 Human pap	179	24.8	88.6	456	10	HPU34108	HPU34108 Human pap
107	28	100.0	7905	10	AF534061	AF534061 Human pap	180	24.8	88.6	456	10	HPU34112	HPU34112
108	28	100.0	7906	10	AY68581	AY68581 Human pap	181	24.8	88.6	456	10	HPU34116	HPU34116
109	28	100.0	7906	10	AY68584	AY68584 Human pap	182	24.8	88.6	456	10	HPU34119	HPU34119
C	110	26.4	149	2	DD187314	DD187314 SPLCBO80	183	24.8	88.6	456	10	HPU34132	HPU34132
111	26.4	94.3	253	10	HPU14511	UI4511 Human pap	184	24.8	88.6	477	10	AF486325	AF486325 Human pap
112	26.4	94.3	273	10	HPU14514	UI4514 Human pap	185	24.8	88.6	543	10	AF404706	AF404706
113	26.4	94.3	298	10	AF404704	AF404704 Human pap	186	24.8	88.6	592	10	AY089923	AY089923 Human pap
114	26.4	94.3	451	10	AF404696	AF404696 Human pap	187	24.8	88.6	602	10	AY112662	AY112662
115	26.4	94.3	451	10	AF404699	AF404699 Human pap	188	24.8	88.6	755	10	AY089952	AY089952 Human pap
116	26.4	94.3	451	10	AF404700	AF404700 Human pap	189	24.8	88.6	1267	10	HPA388065	HPA388065
117	26.4	94.3	451	10	AF404701	AF404701 Human pap	190	24.8	88.6	7906	10	AF402678	AF402678 Human pap
118	26.4	94.3	456	10	AR177943	AR177943 Sequence	191	24.8	88.6	7906	10	AY68579	AY68579 Human pap
119	26.4	94.3	456	10	AF327851	AF327851 Human pap	192	24.8	88.6	7907	10	AY68582	AY68582 Human pap
120	26.4	94.3	456	10	HPU34109	UI4109 Human pap	C	193	78.6	23	2	BD225753	BD225753 Assay uel
121	26.4	94.3	456	10	HPU34110	UI4110 Human pap	C	194	78.6	23	2	AR476279	AR476279 Sequence
122	26.4	94.3	456	10	HPU34113	UI4113 Human pap	C	195	77.9	268478	12	AC182158	AC182158 Bos tauru
123	26.4	94.3	456	10	HPU34114	UI4114 Human pap	C	196	77.9	268478	12	AC182158	AC182158 Bos tauru
124	26.4	94.3	456	10	HPU34117	UI4117 Human pap	C	197	75.7	146976	6	AC171199	AC171199 Mus muscu
125	26.4	94.3	456	10	HPU34118	UI4118 Human pap	C	198	75.7	146976	6	AC173444	AC173444 Atelerix
126	26.4	94.3	456	10	HPU34122	UI4122 Human pap	C	199	75.7	170087	12	AC159465	AC159465 Mus muscu
127	26.4	94.3	456	10	HPU34125	UI4125 Human pap	C	200	75.7	170087	12	AC159465	AC159465 Mus muscu
128	26.4	94.3	456	10	HPU34127	UI4127 Human pap	C	201	75.7	185344	6	AC164163	AC164163 Mus muscu
129	26.4	94.3	456	10	HPU34128	UI4128 Human pap	C	202	75.7	236028	6	CT030235	CT030235
130	26.4	94.3	456	10	HPU34130	UI4130 Human pap	C	203	75.0	92053	12	AC017270	AC017270
131	26.4	94.3	456	10	HPU34131	UI4131 Human pap	C	204	75.0	200859	13	AC010030	AC010030
132	26.4	94.3	456	10	HPU34133	UI4133 Human pap	C	205	75.0	281589	15	AE003546	AE003546
133	26.4	94.3	456	10	HPU34134	UI4134 Human pap	C	206	74.3	149436	12	AC118970	AC118970
134	26.4	94.3	456	10	HPU34135	UI4135 Human pap	C	207	74.3	149436	12	AC103223	AC103223
135	26.4	94.3	456	2	AX658022	AX658022 Sequence	C	208	74.3	227070	12	AC103223	AC103223
136	26.4	94.3	458	10	AF003013	AF003013 Human pap	C	209	74.3	243810	12	AC129394	AC129394
137	26.4	94.3	458	10	AF003014	AF003014 Human pap	C	210	74.3	243810	12	AC154421	AC154421
138	26.4	94.3	458	10	AF003015	AF003015 Human pap	C	211	74.3	247474	12	AC115421	AC115421
139	26.4	94.3	458	10	AF003017	AF003017 Human pap	C	212	73.6	13175	5	AC111193	AC111193 Homo sapi
140	26.4	94.3	458	10	AF003018	AF003018 Human pap	C	213	73.6	13175	5	AC111193	AC111193 Homo sapi
141	26.4	94.3	458	10	AF003019	AF003019 Human pap	C	214	73.6	13175	5	AC111193	AC111193 Homo sapi
142	26.4	94.3	473	10	AF404694	AF404694 Human pap	C	215	73.6	136417	12	AC176718	AC176718 Strongylo
143	26.4	94.3	477	2	CS113240	CS113240 Sequence	C	216	73.6	136417	12	AC176718	AC176718 Strongylo
144	26.4	94.3	477	10	AF486314	AF486314 Human pap	C	217	73.6	141182	12	AC179326	AC179326
145	26.4	94.3	477	10	AF486322	AF486322 Human pap	C	218	73.6	141182	12	AC179326	AC179326
146	26.4	94.3	477	10	AF486332	AF486332 Human pap	C	219	73.6	151188	12	AC180762	AC180762
147	26.4	94.3	477	10	AF486338	AF486338 Human pap	C	220	73.6	151188	12	AC180762	AC180762
148	26.4	94.3	482	10	AF404698	AF404698 Human pap	C	221	73.6	151188	12	AC180762	AC180762
149	26.4	94.3	570	10	AT112663	AT112663 Human pap	C	222	73.6	151188	12	AC180762	AC180762
150	26.4	94.3	600	10	AY098922	AY098922 Human pap	C	223	73.6	169100	6	AC115803	AC115803
151	26.4	94.3	601	10	AY098918	AY098918 Human pap	C	224	73.6	174530	12	AC176694	AC176694
152	26.4	94.3	741	10	AF469197	AF469197 Human pap	C	225	73.6	174530	12	AC176694	AC176694
153	26.4	94.3	755	10	AY089951	AY089951 Human pap	C	226	73.6	182564	12	AC091757	AC091757
154	26.4	94.3	755	10	AY089953	AY089953 Human pap	C	227	73.6	182564	12	AC091757	AC091757
155	26.4	94.3	755	10	AY089954	AY089954 Human pap	C	228	73.6	182564	12	AC091757	AC091757
156	26.4	94.3	755	10	AY089955	AY089955 Human pap	C	229	73.6	206805	12	AC084708	AC084708 Homo sapi
157	26.4	94.3	921	10	HPA388069	HPA388069 Human pap	C	230	73.6	206805	12	AC084708	AC084708 Homo sapi
158	26.4	94.3	1267	10	HPA388057	HPA388057 Human pap	C	231	73.6	206805	12	AC084708	AC084708 Homo sapi
159	26.4	94.3	1267	10	HPA388062	HPA388062 Human pap	C	232	73.6	218754	6	AC125643	AC125643
160	26.4	94.3	1267	10	HPA388063	HPA388063 Human pap	C	233	73.6	218754	6	AC125643	AC125643
161	26.4	94.3	1267	10	HPA388064	HPA388064 Human pap	C	234	73.6	229900	12	AC181347	AC181347
162	26.4	94.3	1267	10	HPA388065	HPA388065 Human pap	C	235	73.6	229900	12	AC181347	AC181347
163	26.4	94.3	1267	10	HPA388067	HPA388067 Human pap	C	236	73.6	25836	12	AC130013	AC130013
164	26.4	94.3	1267	10	HPA388068	HPA388068 Human pap	C	237	73.6	25836	12	AC130013	AC130013
							C	237	73.6	336098	12	AC156054	AC156054

C 238	20.4	72.9	99	2	DD187318	SPLICEOSO	311	19.8	70.7	163576	6	AL807744	AL807744 Mouse DNA
C 239	20.4	72.9	104	2	DD187316	SPLICEOSO	312	19.8	70.7	171142	6	AC113378	AC113378 Rattus no
C 240	20.4	72.9	112591	5	HS189814	AL121966 Human DNA	C 313	19.8	70.7	182088	6	AC100734	AC100734 Mus muscu
C 241	20.4	72.9	141009	12	AC009825	AC009825 Homo sapi	C 314	19.8	70.7	191775	12	AC074358	AC074358 Mus muscu
C 242	20.4	72.9	158765	12	AC178280	AC178280 Strongylo	C 315	19.8	70.7	191833	6	AL161347	AL161347 Mus muscu
C 243	20.4	72.9	170455	12	AC025054	AC025054 Homo sapi	C 316	19.8	70.7	200266	6	AC172315	AC172315 Mouse DNA
C 244	20.4	72.9	211847	12	AC176876	AC176876 Strongylo	C 317	19.8	70.7	205619	4	AC120986	AC120986 Oryza sat
C 245	20.2	72.1	448	2	C0672519	C0672519 Sequence	C 318	19.8	70.7	205616	12	CNS05TDE	AL356033 Homo sapi
C 246	20.2	72.1	42328	6	AL354328	AL354328 Mouse DNA	C 319	19.8	70.7	218519	12	AC135393	AC135393 Rattus no
C 247	20.2	72.1	110000	12	BX248101_2	Continuation (3 of	C 320	19.8	70.7	223246	12	AC164047	AC164047 Bos tauru
C 248	20.2	72.1	110000	15	AE017197_10	Continuation (11 o	C 321	19.8	70.7	223459	12	AC110623	AC110623 Rattus no
C 249	20.2	72.1	112303	5	AC093836	AC093836 Homo sapi	C 322	19.8	70.7	227385	12	AC091407	AC091407 Rattus no
C 250	20.2	72.1	124185	5	AC007560	AC007560 Homo sapi	C 323	19.8	70.7	229035	12	AC117925	AC117925 Rattus no
C 251	20.2	72.1	138006	12	AC068246	AC068246 Homo sapi	C 324	19.8	70.7	246027	12	AC097408	AC097408 Rattus no
C 252	20.2	72.1	153938	12	AC080050	AC080050 Homo sapi	C 325	19.8	70.7	247489	12	AC162989	AC162989 Bos tauru
C 253	20.2	72.1	157978	12	AP001444	AP001444 Homo sapi	C 326	19.8	70.7	249761	12	AC115089	AC115089 Rattus no
C 254	20.2	72.1	162511	5	AC009470	AC009470 Homo sapi	C 327	19.8	70.7	258207	12	AC175556	AC175556 Bos tauru
C 255	20.2	72.1	168738	12	AC019291	AC019291 Homo sapi	C 328	19.6	70.0	423	7	G51331	G51331 SHGC-82851
C 256	20.2	72.1	169470	12	AC019291	AC019291 Homo sapi	C 329	19.6	70.0	751	4	AJ843120	AJ843120 Plantago
C 257	20.2	72.1	170220	5	AC012486	AC012486 Homo sapi	C 330	19.6	70.0	1500	13	AK117096	AK117096 Homo sapi
C 258	20.2	72.1	175872	5	AP001266	AP001266 Homo sapi	C 331	19.6	70.0	4315	4	SBP0A2	X71636 S.bicolor o
C 259	20.2	72.1	196261	5	AC100779	AC100779 Homo sapi	C 332	19.6	70.0	4423	2	C057241	C057241 Sequence
C 260	20.2	72.1	204508	5	AC103833	AC103833 Homo sapi	C 333	19.6	70.0	5240	2	C060764	C060764 Sequence
C 261	20.2	72.1	209242	5	CNS01RGM	AL157996 Human chr	C 334	19.6	70.0	7082	12	AC014876	AC014876 Zebrafish
C 262	20.2	72.1	220028	12	AC153442	AC153442 Bos tauru	C 335	19.6	70.0	28809	11	BX539306	BX539306 Zebrafish
C 263	20.2	72.1	232443	12	AC099251	AC099251 Rattus no	C 336	19.6	70.0	39472	5	HSN7A10	Z68324 Human DNA s
C 264	20	71.4	2135	11	BC108531	BC108531 Xenopus l	C 337	19.6	70.0	39624	5	AC069504	AC069504 Homo sapi
C 265	20	71.4	33398	4	SPBC388	AL022244 S.pombe c	C 338	19.6	70.0	41449	6	AC167190	AC167190 Mus muscu
C 266	20	71.4	80472	4	F24U2	AF262039 Arabidops	C 339	19.6	70.0	43115	12	AC162070	AC162070 Bos tauru
C 267	20	71.4	84544	4	AC007134	AC007134 Arabidops	C 340	19.6	70.0	49162	12	AC160518	AC160518 Strongylo
C 268	20	71.4	87119	4	AC005897	AC005897 Arabidops	C 341	19.6	70.0	62999	12	AC087594	AC087594 Homo sapi
C 269	20	71.4	88239	12	AP007978	AP007978 Arabidops	C 342	19.6	70.0	66608	12	AC162237	AC162237 Bos tauru
C 270	20	71.4	97864	4	AP006093	AP006093 Lotus cor	C 343	19.6	70.0	72208	12	AC178420	AC178420 Strongylo
C 271	20	71.4	104684	4	AP006124	AP006124 Lotus cor	C 344	19.6	70.0	74129	12	AC161952	AC161952 Bos tauru
C 272	20	71.4	117945	4	AC007045	AC007045 Arabidops	C 345	19.6	70.0	90971	12	AC169058	AC169058 Bos tauru
C 273	20	71.4	135391	11	BX465182	BX465183 Zebrafish	C 346	19.6	70.0	94348	11	AC109580	AC109580 Dario rer
C 274	20	71.4	135391	5	HS805C22	AL035667 Human DNA	C 347	19.6	70.0	98558	5	AL137138	AL137138 Human DNA
C 275	20	71.4	137336	4	AC007915	AC007915 Genomic s	C 348	19.6	70.0	101325	12	AC166898	AC166898 Medicago
C 276	20	71.4	147738	12	CR848712	CR848712 Dario rer	C 349	19.6	70.0	103541	12	AC158508	AC158508 Pan trogl
C 277	20	71.4	148870	5	HS173A13	AL035668 Human DNA	C 350	19.6	70.0	105001	2	ARE59586	ARE59586 Sequence
C 278	20	71.4	159852	5	AC097476	AC097476 Homo sapi	C 351	19.6	70.0	105002	2	ARE59780	ARE59780 Sequence
C 279	20	71.4	170095	5	AL356131	AL356131 Human DNA	C 352	19.6	70.0	106321	12	AC148344	AC148344 Medicago
C 280	20	71.4	175008	12	AC025879	AC025879 Homo sapi	C 353	19.6	70.0	106951	4	AC147963	AC147963 Medicago
C 281	20	71.4	176009	11	BX465834	BX465834 Zebrafish	C 354	19.6	70.0	110000	4	AP008207_358	Continuation (359
C 282	20	71.4	181535	6	AC123529	AC123529 Mus muscu	C 355	19.6	70.0	110000	4	AP008209_322	Continuation (332
C 283	20	71.4	183355	12	AC1178035	AC1178035 Strongylo	C 356	19.6	70.0	110000	15	AE017321_00	Continuation (2 of
C 284	20	71.4	225738	6	AC102633	AC102633 Mus muscu	C 357	19.6	70.0	110000	15	AE017321_01	Continuation (2 of
C 285	20	71.4	230139	6	AC116731	AC116731 Mus muscu	C 358	19.6	70.0	113259	12	AC181952	AC181952 Strongylo
C 286	20	71.4	255663	6	AC091276	AC091276 Mus muscu	C 359	19.6	70.0	121026	6	BX545905	AC146777 Medicago
C 287	20	71.4	273739	12	CR848028	CR848028 Dario rer	C 360	19.6	70.0	128574	5	AC004080	AC004080 Homo sapi
C 288	20	71.4	278729	12	AC106916	AC106916 Rattus no	C 361	19.6	70.0	129355	4	AC091123	AC091123 Oryza sat
C 289	20	71.4	300661	12	AC106916	AC106916 Rattus no	C 362	19.6	70.0	129778	4	AC091123	AC091123 Oryza sat
C 290	19.8	70.7	4864	15	AJ583758	AJ583758 Haemophil	C 363	19.6	70.0	134355	12	AC176911	AC176911 Strongylo
C 291	19.8	70.7	4623	15	AJ937360	AJ937360 Haemophil	C 364	19.6	70.0	142991	12	AC151805	AC151805 Medicago
C 292	19.8	70.7	22993	2	DD194608	DD194608 Method an	C 365	19.6	70.0	144047	12	AC175222	AC175222 Atelerix
C 293	19.8	70.7	22993	2	DD194608	DD194608 Method an	C 366	19.6	70.0	152113	5	AL592285	AL592285 Human DNA
C 294	19.8	70.7	22993	2	AX825168	AX825168 Sequence	C 367	19.6	70.0	154772	13	AC009214	AC009214 Drosophil
C 295	19.8	70.7	22993	2	AX825170	AX825170 Sequence	C 368	19.6	70.0	156381	11	AL954313	AL954313 Zebrafish
C 296	19.8	70.7	38342	2	AX251503	AX251503 Sequence	C 369	19.6	70.0	157544	12	AC162255	AC162255 Bos tauru
C 297	19.8	70.7	38342	2	AX344502	AX344502 Sequence	C 370	19.6	70.0	159584	12	AC150232	AC150232 Papio anu
C 298	19.8	70.7	87937	5	CNS05TDEA	AL583373 Human chr	C 371	19.6	70.0	160190	12	AC158833	AC158833 Bos tauru
C 299	19.8	70.7	89995	5	AC133794	AC133794 Homo sapi	C 372	19.6	70.0	161134	5	AC079161	AC079161 Homo sapi
C 300	19.8	70.7	110000	4	AP008211_1280	Continuation (13 o	C 373	19.6	70.0	161654	5	AC116608	AC116608 Papio ham
C 301	19.8	70.7	110000	4	AP008211_1280	Continuation (13 o	C 374	19.6	70.0	162215	6	AL731773	AL731773 Mouse DNA
C 302	19.8	70.7	110000	13	CR548612_00	CR548612 Paramectu	C 375	19.6	70.0	162996	12	AL391359	AL391359 Homo sapi
C 303	19.8	70.7	149167	12	AC158838	AC158838 Bos tauru	C 376	19.6	70.0	163050	12	AC020696	AC020696 Homo sapi
C 304	19.8	70.7	149358	6	AC087251	AC087251 Rattus no	C 377	19.6	70.0	163640	5	AC093328	AC093328 Homo sapi
C 305	19.8	70.7	150568	12	AC168579	AC168579 Strongylo	C 378	19.6	70.0	165194	12	AC176112	AC176112 Strongylo
C 306	19.8	70.7	156196	11	BX072559	BX072559 Zebrafish	C 379	19.6	70.0	165194	12	AC149184	AC149184 Papio anu
C 307	19.8	70.7	157916	11	BX511133	BX511133 Zebrafish	C 380	19.6	70.0	167794	12	AP215844	AP215844 Homo sapi
C 308	19.8	70.7	159933	12	AC170027	AC170027 Bos tauru	C 381	19.6	70.0	169511	13	AC007531	AC007531 Drosophil
C 309	19.8	70.7	160907	4	AC120990	AC120990 Oryza sat	C 382	19.6	70.0	170193	11	AL935189	AL935189 Zebrafish
C 310	19.8	70.7	161067	5	AL353638	AL353638 Human DNA	C 383	19.6	70.0	170759	4	AP003271	AP003271 Oryza sat

C 384	19.6	70.0	171162	12	AC068684	Homo sapi	457	19.2	68.6	118466	12	AC178189	Strongylo	457
C 385	19.6	70.0	177556	12	AC163233	Bos tauru	458	19.2	68.6	121896	5	AL138713	Human DNA	458
C 386	19.6	70.0	183474	12	AC153103	Sus scrofa	C 459	19.2	68.6	123427	4	AC091732	Oryza sat	C 459
C 387	19.6	70.0	184498	12	AC173589	Bos tauru	460	19.2	68.6	125573	12	AC010112	Drosophila	460
C 388	19.6	70.0	186466	5	AC054475	Homo sapi	461	19.2	68.6	125742	12	AC018176	Strongylo	461
C 389	19.6	70.0	187825	12	AC161061	Pan trogl	462	19.2	68.6	126420	12	AC179129	Strongylo	462
C 390	19.6	70.0	193390	6	AC101527	Mus muscu	C 463	19.2	68.6	128777	12	AC017835	Triticum	C 463
C 391	19.6	70.0	193641	6	AC121972	Mus muscu	C 464	19.2	68.6	132655	13	AC010015	Drosophila	C 464
C 392	19.6	70.0	193670	5	AC026261	Mus muscu	C 465	19.2	68.6	134430	12	AC151300	Strongylo	C 465
C 393	19.6	70.0	200397	5	AC026261	Mus muscu	C 466	19.2	68.6	135291	12	AC178871	Strongylo	C 466
C 394	19.6	70.0	200997	12	AC144284	Homo sapi	C 467	19.2	68.6	136846	12	AC179413	Strongylo	C 467
C 395	19.6	70.0	201854	5	AC098934	Mus muscu	C 468	19.2	68.6	137211	12	AC180440	Strongylo	C 468
C 396	19.6	70.0	210320	5	AC093873	Homo sapi	C 469	19.2	68.6	140030	12	AC152550	Xenopus t	C 469
C 397	19.6	70.0	212257	12	AC168612	Strongylo	C 470	19.2	68.6	142695	6	AC163386	Mus muscu	C 470
C 398	19.6	70.0	212279	12	AC168612	Strongylo	C 471	19.2	68.6	144015	12	AC181823	Strongylo	C 471
C 399	19.6	70.0	212749	12	AC115435	Rattus no	C 472	19.2	68.6	145373	11	BX649288	Zebrafish	C 472
C 400	19.6	70.0	213045	12	AC170084	Rattus no	C 473	19.2	68.6	147616	12	AC176425	Strongylo	C 473
C 401	19.6	70.0	221023	12	AC174692	Medicago	C 474	19.2	68.6	148103	12	AC177109	Strongylo	C 474
C 402	19.6	70.0	222328	12	AC172864	Bos tauru	C 475	19.2	68.6	148704	4	AC124213	Genomic s	C 475
C 403	19.6	70.0	222328	12	AC175761	Bos tauru	C 476	19.2	68.6	148924	12	AC161141	Bos tauru	C 476
C 404	19.6	70.0	228978	12	AC094027	Rattus no	C 477	19.2	68.6	151102	12	AC177554	Strongylo	C 477
C 405	19.6	70.0	232952	12	AC120963	Rattus no	C 478	19.2	68.6	151190	12	BX640543	Danio rer	C 478
C 406	19.6	70.0	233874	12	BX845298	BX845298	C 479	19.2	68.6	152368	12	BX640541	Danio rer	C 479
C 407	19.6	70.0	234470	11	BX539340	Zebrafish	C 480	19.2	68.6	153652	12	AC175063	Bos tauru	C 480
C 408	19.6	70.0	239863	12	AC106948	Rattus no	C 481	19.2	68.6	154018	12	AC065019	Homo sapi	C 481
C 409	19.6	70.0	244833	12	AC169590	Bos tauru	C 482	19.2	68.6	156083	5	AC074087	Homo sapi	C 482
C 410	19.6	70.0	246231	13	AB003636	Rattus no	C 483	19.2	68.6	156937	12	AC176805	Strongylo	C 483
C 411	19.6	70.0	248776	12	AC115447	Rattus no	C 484	19.2	68.6	158391	12	AC008023	Strongylo	C 484
C 412	19.6	70.0	250968	12	AC014140	Drosophila	C 485	19.2	68.6	158669	12	AC074238	Strongylo	C 485
C 413	19.6	70.0	254464	12	AC011740	Rattus no	C 486	19.2	68.6	158736	12	AC180732	Strongylo	C 486
C 414	19.6	70.0	262915	12	AC170788	Bos tauru	C 487	19.2	68.6	159035	5	AC092584	Homo sapi	C 487
C 415	19.6	70.0	262915	12	AC170788	Bos tauru	C 488	19.2	68.6	159150	12	AC1172408	Strongylo	C 488
C 416	19.6	70.0	286083	12	AC171475	Bos tauru	C 489	19.2	68.6	159264	5	AC007653	Homo sapi	C 489
C 417	19.6	70.0	299749	13	AB003606	Drosophila	C 490	19.2	68.6	160214	6	AL928566	Mouse DNA	C 490
C 418	19.4	69.3	70985	12	AC175758	Bos tauru	C 491	19.2	68.6	160828	12	AC181217	Strongylo	C 491
C 419	19.4	69.3	110000	12	AC127655_4	Continuation (5 of	C 492	19.2	68.6	161300	12	AC026169	Homo sapi	C 492
C 420	19.4	69.3	127439	5	AC126178	Continuation (2 of	C 493	19.2	68.6	161551	12	AC180899	Strongylo	C 493
C 421	19.4	69.3	172239	6	AC126679	AC126679 Homo sapi	C 494	19.2	68.6	162497	12	AC022172	Homo sapi	C 494
C 422	19.4	69.3	182043	14	AC150593	AC150593 Bos tauru	C 495	19.2	68.6	162942	12	AC180189	Strongylo	C 495
C 423	19.4	69.3	203839	6	AC161265	AC161265 Mus muscu	C 496	19.2	68.6	164310	12	AC181652	Strongylo	C 496
C 424	19.4	69.3	210741	6	AC162139	AC162139 Mus muscu	C 497	19.2	68.6	165318	13	AC010061	Drosophila	C 497
C 425	19.4	69.3	229771	12	AC137170	AC137170 Rattus no	C 498	19.2	68.6	165993	11	AL772362	Zebrafish	C 498
C 426	19.4	69.3	337203	13	CR382401	CR382401 Plasmodiu	C 499	19.2	68.6	166253	12	AC175783	Triticum	C 499
C 427	19.2	68.6	348	2	AX088033	AX088033 Sequence	C 500	19.2	68.6	166689	12	AC058815	Homo sapi	C 500
C 428	19.2	68.6	522	2	AX366832	AX366832 Sequence	C 501	19.2	68.6	167152	5	AC092839	Homo sapi	C 501
C 429	19.2	68.6	745	11	GGMHG4A6	X63087 G.gallus HM	C 502	19.2	68.6	167507	12	AC176021	Strongylo	C 502
C 430	19.2	68.6	872	11	CHKMG14A	M26675 Gallus gall	C 503	19.2	68.6	167927	12	AC162782	Homo sapi	C 503
C 431	19.2	68.6	2102	4	AF147497	AF147497 Humulus l	C 504	19.2	68.6	169067	12	AC016287	Homo sapi	C 504
C 432	19.2	68.6	4010	2	AX088031	AX088031 Sequence	C 505	19.2	68.6	170119	12	AC176866	Strongylo	C 505
C 433	19.2	68.6	5536	2	AX345081	AX345081 Sequence	C 506	19.2	68.6	170654	6	AC116502	Homo sapi	C 506
C 434	19.2	68.6	6389	2	AX346863	AX346863 Sequence	C 507	19.2	68.6	171905	12	AL591609	Homo sapi	C 507
C 435	19.2	68.6	23448	5	AL355174	AL355174 Human DNA	C 508	19.2	68.6	173210	11	AL928968	Homo sapi	C 508
C 436	19.2	68.6	30220	12	AC019816	AC019816 Drosophila	C 509	19.2	68.6	173358	13	AC150550	Drosophila	C 509
C 437	19.2	68.6	37887	12	AC068453	AC068453 Homo sapi	C 510	19.2	68.6	174390	11	BX323991	Drosophila	C 510
C 438	19.2	68.6	54767	11	AC157739	AC157739 Xenopus t	C 511	19.2	68.6	174939	13	AC008256	Drosophila	C 511
C 439	19.2	68.6	71197	11	CR848758	CR848758 Zebrafish	C 512	19.2	68.6	176730	12	AC073994	Homo sapi	C 512
C 440	19.2	68.6	82632	12	AC008181	AC008181 Drosophila	C 513	19.2	68.6	177644	5	AC023442	Homo sapi	C 513
C 441	19.2	68.6	82632	12	AC008181	AC008181 Drosophila	C 514	19.2	68.6	178351	5	AC099542	Homo sapi	C 514
C 442	19.2	68.6	100128	11	AC020135	AC020135 Xenopus t	C 515	19.2	68.6	179321	6	AC101719	Mus muscu	C 515
C 443	19.2	68.6	102986	12	AC020135	AC020135 Drosophila	C 516	19.2	68.6	181056	13	AC009733	Drosophila	C 516
C 444	19.2	68.6	103980	5	AL354893	AL354893 Human DNA	C 517	19.2	68.6	181079	12	AC016667	Homo sapi	C 517
C 445	19.2	68.6	104563	12	AC180038	AC180038 Strongylo	C 518	19.2	68.6	181925	12	AC020710	Homo sapi	C 518
C 446	19.2	68.6	107857	5	AC008631	AC008631 Homo sapi	C 519	19.2	68.6	183020	13	AC015056	Homo sapi	C 519
C 447	19.2	68.6	108482	12	AC180747	AC180747 Strongylo	C 520	19.2	68.6	183596	5	AC008172	Homo sapi	C 520
C 448	19.2	68.6	110000	4	AP008215_119	Continuation (120	C 521	19.2	68.6	184566	6	AC123827	Mus muscu	C 521
C 449	19.2	68.6	110000	4	CR382125_11	Continuation (120	C 522	19.2	68.6	185363	6	AC123827	Mus muscu	C 522
C 450	19.2	68.6	110000	4	AE016955_119	Continuation (120	C 523	19.2	68.6	185655	12	AC179192	Strongylo	C 523
C 451	19.2	68.6	110000	12	AC116854_1	Continuation (15 o	C 524	19.2	68.6	187778	12	AC027180	Homo sapi	C 524
C 452	19.2	68.6	113000	12	TANN3_14	Continuation (15 o	C 525	19.2	68.6	187793	13	AC013247	Drosophila	C 525
C 453	19.2	68.6	113367	5	AC008178	AC008178 Homo sapi	C 526	19.2	68.6	193399	12	AC116541	Drosophila	C 526
C 454	19.2	68.6	115272	11	BX511080	BX511080 Zebrafish	C 527	19.2	68.6	197864	5	AC021956	Homo sapi	C 527
C 455	19.2	68.6	117939	12	CT573051	CT573051 Medicago	C 528	19.2	68.6	198217	5	AL354942	Human DNA	C 528
C 456	19.2	68.6	118151	5	HS249C1	AL022154 Human DNA	C 529	19.2	68.6	199852	6	AC163746	Mus muscu	C 529

530	19.2	68.6	202311	12	AC179831	Strongylo	603	19	67.9	1981	2	AX764575	Sequence
C 531	19.2	68.6	205873	12	AC073147	Homo sapi	C 604	19	67.9	2119	6	BC101869	Strongylo
C 532	19.2	68.6	209293	6	AC153906	Mus muscu	C 605	19	67.9	2834	4	CPU69275	Cryptonec
C 533	19.2	68.6	210643	12	AC103432	Rattus no	C 606	19	67.9	2901	6	AI591743	Mus muscu
C 534	19.2	68.6	212426	5	AC104335	Homo sapi	C 607	19	67.9	6092	12	AX345922	Sequence
C 535	19.2	68.6	216157	12	APF485269	Mus muscu	C 608	19	67.9	12231	15	AE002118	Ureaplas
C 536	19.2	68.6	219254	12	AC172524	Bos tauru	C 609	19	67.9	12415	13	CEY102F5A	Caenorhab
C 537	19.2	68.6	221989	5	AY220878	Homo sapi	C 610	19	67.9	17684	14	RBA511CG	Strongylo
C 538	19.2	68.6	228466	12	AC170217	Bos tauru	C 611	19	67.9	18008	5	AC093147	Homo sapi
C 539	19.2	68.6	228443	13	AE003772	Drosophi	C 612	19	67.9	20112	14	AI284844	Orycto
C 540	19.2	68.6	229087	12	AC181100	Strongylo	C 613	19	67.9	30082	5	AI606759	Human DNA
C 541	19.2	68.6	230267	11	BX324146	Zebrafish	C 614	19	67.9	33745	13	U97551	Caenorhabd
C 542	19.2	68.6	237800	12	AC172126	Bos tauru	C 615	19	67.9	34009	5	AC105447	Homo sapi
C 543	19.2	68.6	240562	12	AC166667	Bos tauru	C 616	19	67.9	42301	5	AP000528	Homo sapi
C 544	19.2	68.6	243188	12	AC106625	Rattus no	C 617	19	67.9	43503	12	AC150000	Strongylo
C 545	19.2	68.6	243188	12	AC106625	Rattus no	C 618	19	67.9	43503	12	AC150000	Strongylo
C 546	19.2	68.6	244237	6	AC136921	Mus muscu	C 619	19	67.9	55012	5	AL136976	Human DNA
C 547	19.2	68.6	245819	12	AC172437	Bos tauru	C 620	19	67.9	55722	12	AC168987	Bos tauru
C 548	19.2	68.6	250036	12	AC125669	Rattus no	C 621	19	67.9	55981	11	CR388122	Zebrafish
C 549	19.2	68.6	252928	12	AC166438	Bos tauru	C 622	19	67.9	56448	13	AC006679	Caenorhab
C 550	19.2	68.6	253434	12	AC170235	Bos tauru	C 623	19	67.9	64976	12	AC026256	Homo sapi
C 551	19.2	68.6	253950	12	AC111282	Rattus no	C 624	19	67.9	65556	12	AC100220	Mus muscu
C 552	19.2	68.6	254107	12	AC169777	Bos tauru	C 625	19	67.9	66252	12	AC108772	Mus muscu
C 553	19.2	68.6	265308	12	AC173276	Bos tauru	C 626	19	67.9	67644	12	AC100474	Mus muscu
C 554	19.2	68.6	268549	12	AC136581	Rattus no	C 627	19	67.9	67644	12	AC100474	Mus muscu
C 555	19.2	68.6	271869	13	AE003531	Drosophi	C 628	19	67.9	72371	12	AC113530	Mus muscu
C 556	19.2	68.6	272890	12	AC016653	Homo sapi	C 629	19	67.9	78928	11	AC119990	Strongylo
C 557	19.2	68.6	294609	13	AE003787	Drosophi	C 630	19	67.9	79383	4	CNS08CDQ	Strongylo
C 558	19.2	68.6	301146	12	AC111317	Rattus no	C 631	19	67.9	95337	5	AB005231	Atribidops
C 559	19.2	68.6	349980	2	AX344555	Sequence	C 632	19	67.9	96514	5	AL136117	Human DNA
C 560	19.2	68.6	349980	2	AX344555	Sequence	C 633	19	67.9	90139	12	AC174600	Strongylo
C 561	19	67.9	305	2	AX363421	Sequence	C 634	19	67.9	91187	5	AC023790	Strongylo
C 562	19	67.9	438	12	CQ425759	Sequence	C 635	19	67.9	92202	12	AC168547	Strongylo
C 563	19	67.9	451	13	AB106560	Tubercula	C 636	19	67.9	94301	4	ATF28A21	Strongylo
C 564	19	67.9	600	7	BV242113	BV242113	C 637	19	67.9	95377	12	CT573500	Strongylo
C 565	19	67.9	613	7	BV217366	BV217366	C 638	19	67.9	99226	12	AC178812	Strongylo
C 566	19	67.9	650	7	BV063711	BV063711	C 639	19	67.9	99940	5	AC079246	Strongylo
C 567	19	67.9	766	7	BV570846	BV570846	C 640	19	67.9	100122	12	AP178220	Strongylo
C 568	19	67.9	811	7	BV044098	BV044098	C 641	19	67.9	100709	5	AC073319	Strongylo
C 569	19	67.9	819	7	BV557936	BV557936	C 642	19	67.9	103090	4	AC152184	Strongylo
C 570	19	67.9	853	7	BV560472	BV560472	C 643	19	67.9	110000	4	AP008214_028	Strongylo
C 571	19	67.9	854	7	BV560472	BV560472	C 644	19	67.9	110000	4	AP008214_008	Strongylo
C 572	19	67.9	858	7	BV568956	BV568956	C 645	19	67.9	110000	4	CR382129_21	Strongylo
C 573	19	67.9	869	7	AK063812	Oryza sat	C 646	19	67.9	110000	4	AP008207_030	Strongylo
C 574	19	67.9	944	4	AK102322	Oryza sat	C 647	19	67.9	110000	12	AC109862_1	Strongylo
C 575	19	67.9	1275	8	AY774156	Synthetic	C 648	19	67.9	110000	12	AY657022_3	Strongylo
C 576	19	67.9	1386	8	AY872171	Synthetic	C 649	19	67.9	110000	12	AY657022_3	Strongylo
C 577	19	67.9	1419	15	FX15789	Peridobact	C 650	19	67.9	110000	15	AB017261_03	Strongylo
C 578	19	67.9	1492	4	PN1580714	PN1580714	C 651	19	67.9	110000	15	AB017261_03	Strongylo
C 579	19	67.9	1492	4	PN1580715	PN1580715	C 652	19	67.9	110000	15	AB017261_03	Strongylo
C 580	19	67.9	1492	4	PN1580716	PN1580716	C 653	19	67.9	110520	5	AL603757	Human DNA
C 581	19	67.9	1492	4	PN1580717	PN1580717	C 654	19	67.9	111786	4	AC152351	Human DNA
C 582	19	67.9	1492	4	PN1580718	PN1580718	C 655	19	67.9	111786	4	AC152351	Human DNA
C 583	19	67.9	1492	4	PN1580719	PN1580719	C 656	19	67.9	111786	4	AC152351	Human DNA
C 584	19	67.9	1492	4	PN1580720	PN1580720	C 657	19	67.9	111786	4	AC152351	Human DNA
C 585	19	67.9	1492	4	PN1580721	PN1580721	C 658	19	67.9	111786	4	AC152351	Human DNA
C 586	19	67.9	1492	4	PN1580722	PN1580722	C 659	19	67.9	111786	4	AC152351	Human DNA
C 587	19	67.9	1492	4	PN1580723	PN1580723	C 660	19	67.9	111786	4	AC152351	Human DNA
C 588	19	67.9	1492	4	PN1580724	PN1580724	C 661	19	67.9	111786	4	AC152351	Human DNA
C 589	19	67.9	1493	4	PN1580705	PN1580705	C 662	19	67.9	120450	5	AC018456	Strongylo
C 590	19	67.9	1493	4	PN1580706	PN1580706	C 663	19	67.9	120927	5	AC092575	Strongylo
C 591	19	67.9	1493	4	PN1580707	PN1580707	C 664	19	67.9	120927	5	AC092575	Strongylo
C 592	19	67.9	1493	4	PN1580708	PN1580708	C 665	19	67.9	120927	5	AC092575	Strongylo
C 593	19	67.9	1493	4	PN1580709	PN1580709	C 666	19	67.9	120927	5	AC092575	Strongylo
C 594	19	67.9	1493	4	PN1580710	PN1580710	C 667	19	67.9	120927	5	AC092575	Strongylo
C 595	19	67.9	1493	4	PN1580711	PN1580711	C 668	19	67.9	120927	5	AC092575	Strongylo
C 596	19	67.9	1493	4	PN1580712	PN1580712	C 669	19	67.9	120927	5	AC092575	Strongylo
C 597	19	67.9	1493	4	PN1580713	PN1580713	C 670	19	67.9	120927	5	AC092575	Strongylo
C 598	19	67.9	1493	4	PN1580714	PN1580714	C 671	19	67.9	120927	5	AC092575	Strongylo
C 599	19	67.9	1493	4	PN1580715	PN1580715	C 672	19	67.9	120927	5	AC092575	Strongylo
C 600	19	67.9	1493	4	PN1580716	PN1580716	C 673	19	67.9	120927	5	AC092575	Strongylo
C 601	19	67.9	1493	4	PN1580717	PN1580717	C 674	19	67.9	120927	5	AC092575	Strongylo
C 602	19	67.9	1493	4	PN1580718	PN1580718	C 675	19	67.9	120927	5	AC092575	Strongylo

676	19	67.9	141672	5	HS142F18	AL031073 Human DNA	c 749	19	67.9	184284	12	AC068054	AC068054 Homo sapi
677	19	67.9	141939	5	AP466646	AP466646 Zee maye	c 750	19	67.9	184328	12	AC136380	AC136380 Rattus no
678	19	67.9	142557	5	AL356499	AL356499 Human DNA	c 751	19	67.9	184770	5	AP005670	AP005670 Homo sapi
679	19	67.9	143113	12	AC176478	AC176478 Strongylo	c 752	19	67.9	186416	12	AC011226	AC011226 Homo sapi
680	19	67.9	144136	6	AC122814	AC122814 Mus muscu	c 753	19	67.9	186448	12	AC034129	AC034129 Homo sapi
681	19	67.9	144172	6	AC154745	AC154745 Mus muscu	c 754	19	67.9	187674	12	AC013748	AC013748 Homo sapi
682	19	67.9	145246	6	AC115906	AC115906 Mus muscu	c 755	19	67.9	187981	12	AC013786	AC013786 Homo sapi
683	19	67.9	145451	4	AP005824	AP005824 Pan trogl	c 756	19	67.9	188000	6	AC160053	AC160053 Mus muscu
684	19	67.9	146335	12	AP002521	AP002521 Oryza sat	c 757	19	67.9	189088	12	AC146672	AC146672 Oryza sat
685	19	67.9	146897	12	CR457454	CR457454 Dario rer	c 758	19	67.9	189276	5	AP006261	AP006261 Homo sapi
686	19	67.9	148997	5	AC021015	AC021015 Homo sapi	c 759	19	67.9	190294	12	AC160285	AC160285 Bos tauru
687	19	67.9	150022	12	AC168794	AC168794 Strongylo	c 760	19	67.9	190369	6	AC119908	AC119908 Mus muscu
688	19	67.9	150150	4	AP002539	AP002539 Oryza sat	c 761	19	67.9	190621	6	AC161451	AC161451 Mus muscu
689	19	67.9	151636	12	AC180903	AC180903 Strongylo	c 762	19	67.9	191105	12	AC150473	AC150473 Papio anu
690	19	67.9	151698	12	AC132644	AC132644 Rattus no	c 763	19	67.9	191635	6	BX119996	BX119996 Zebrafish
691	19	67.9	151970	12	AC040905	AC040905 Homo sapi	c 764	19	67.9	192849	11	BX572079	BX572079 Zebrafish
692	19	67.9	152058	4	AP005740	AP005740 Oryza sat	c 765	19	67.9	193120	12	AC113559	AC113559 Canis fam
693	19	67.9	152939	12	AC178269	AC178269 Strongylo	c 766	19	67.9	193250	12	AC165173	AC165173 Mus muscu
694	19	67.9	153099	12	AC149983	AC149983 Strongylo	c 767	19	67.9	193420	11	CR513789	CR513789 Strongylo
695	19	67.9	153878	12	AC159170	AC159170 Papio anu	c 768	19	67.9	196536	12	AC180373	AC180373 Strongylo
696	19	67.9	154078	5	CNS07EFD	AL512624 Human chr	c 769	19	67.9	196737	6	CT025551	CT025551 Mouse DNA
697	19	67.9	158469	5	AC012082	AC012082 Homo sapi	c 770	19	67.9	199075	4	ATC8R1V49	ATC8R1V49 Homo sapi
698	19	67.9	158857	12	AC023522	AC023522 Homo sapi	c 771	19	67.9	201733	12	AC068851	AC068851 Bos tauru
699	19	67.9	160187	12	AC178605	AC178605 Strongylo	c 772	19	67.9	203651	12	AC160158	AC160158 Homo sapi
700	19	67.9	161840	5	CNS07BG2	AL589182 Human chr	c 773	19	67.9	203810	6	AC132612	AC132612 Mus muscu
701	19	67.9	162090	12	AC147537	AC147537 Medicago	c 774	19	67.9	203877	6	AC122260	AC122260 Mus muscu
702	19	67.9	162139	4	AP004461	AP004461 Oryza sat	c 775	19	67.9	204800	6	AC132126	AC132126 Zebrafish
703	19	67.9	162223	11	AL845363	AL845363 Zebrafish	c 776	19	67.9	205347	11	BX649534	BX649534 Homo sapi
704	19	67.9	162935	11	CR847843	CR847843 Zebrafish	c 777	19	67.9	206153	12	AC170914	AC170914 Mus muscu
705	19	67.9	163698	6	AC115876	AC115876 Mus muscu	c 778	19	67.9	208153	6	AC158364	AC158364 Rattus no
706	19	67.9	163976	6	AC154672	AC154672 Mus muscu	c 779	19	67.9	208726	12	AC134638	AC134638 Bos tauru
707	19	67.9	163976	12	AC080000	AC080000 Homo sapi	c 780	19	67.9	211100	12	AC170572	AC170572 Bos tauru
708	19	67.9	164065	14	AC147455	AC147455 Atelestix	c 781	19	67.9	211775	12	AC027205	AC027205 Homo sapi
709	19	67.9	164239	5	EX072566	EX072566 Human DNA	c 782	19	67.9	211812	12	AC027205	AC027205 Homo sapi
710	19	67.9	164595	5	AC024560	AC024560 Homo sapi	c 783	19	67.9	213497	12	AL591966	AL591966 Bos tauru
711	19	67.9	164615	12	AC152218	AC152218 Bos tauru	c 784	19	67.9	214053	12	AC161290	AC161290 Rattus no
712	19	67.9	164628	5	CNS01DVJ	AL135858 Human chr	c 785	19	67.9	215917	12	AC112385	AC112385 Bos tauru
713	19	67.9	165767	6	AC114918	AC114918 Mus muscu	c 786	19	67.9	215928	12	AC166697	AC166697 Bos tauru
714	19	67.9	166125	12	AL365448	AL365448 Homo sapi	c 787	19	67.9	215951	12	AC170940	AC170940 Bos tauru
715	19	67.9	166197	6	AC132343	AC132343 Mus muscu	c 788	19	67.9	219166	6	AC152552	AC152552 Medicago
716	19	67.9	166239	5	HS214614	AL050303 Homo sapi	c 789	19	67.9	222939	6	AC124568	AC124568 Homo sapi
717	19	67.9	166440	12	AC149747	AC149747 Bos tauru	c 790	19	67.9	222946	12	AC021978	AC021978 Bos tauru
718	19	67.9	166929	5	AC084730	AC084730 Papio anu	c 791	19	67.9	223797	12	AC021279	AC021279 Homo sapi
719	19	67.9	168097	12	AC177377	AC177377 Strongylo	c 792	19	67.9	223985	12	AC130648	AC130648 Rattus no
720	19	67.9	168638	5	AC009433	AC009433 Homo sapi	c 793	19	67.9	224792	12	AC130648	AC130648 Rattus no
721	19	67.9	171805	12	AC087607	AC087607 Homo sapi	c 794	19	67.9	225445	12	AC134035	AC134035 Rattus no
722	19	67.9	172237	5	AC139112	AC139112 Cercopithec	c 795	19	67.9	225630	12	AC134035	AC134035 Rattus no
723	19	67.9	172545	5	AL162851	AL162851 Human DNA	c 796	19	67.9	226052	6	AC099690	AC099690 Mus muscu
724	19	67.9	172751	12	CT030368	CT030368 Platyrrhine	c 797	19	67.9	227241	11	AC111995	AC111995 Homo sapi
725	19	67.9	173048	5	AC096538	AC096538 Homo sapi	c 798	19	67.9	227616	11	BX072555	BX072555 Zebrafish
726	19	67.9	173048	6	AC115878	AC115878 Mus muscu	c 799	19	67.9	227632	5	AC087382	AC087382 Homo sapi
727	19	67.9	173458	5	AC092136	AC092136 Homo sapi	c 800	19	67.9	227709	12	AC112288	AC112288 Homo sapi
728	19	67.9	173548	5	AC092136	AC092136 Homo sapi	c 801	19	67.9	227709	12	AC079421	AC079421 Homo sapi
729	19	67.9	173781	12	AC181149	AC181149 Strongylo	c 802	19	67.9	227958	12	BX897723	BX897723 Homo sapi
730	19	67.9	173905	5	AC092125	AC092125 Homo sapi	c 803	19	67.9	228377	12	AC117358	AC117358 Rattus no
731	19	67.9	175789	12	AP000425	AP000425 Homo sapi	c 804	19	67.9	228986	12	AC117358	AC117358 Rattus no
732	19	67.9	176156	5	CNS08CCG	AL292961 Human chr	c 805	19	67.9	229363	11	BX324135	BX324135 Rattus no
733	19	67.9	176816	12	AC113472	AC113472 Mus muscu	c 806	19	67.9	231236	12	AC096369	AC096369 Rattus no
734	19	67.9	178029	5	AC093838	AC093838 Homo sapi	c 807	19	67.9	231236	12	AC096369	AC096369 Rattus no
735	19	67.9	178563	11	AL928650	AL928650 Zebrafish	c 808	19	67.9	231915	11	BX649405	BX649405 Homo sapi
736	19	67.9	178862	5	AC011847	AC011847 Homo sapi	c 809	19	67.9	232525	5	AC127023	AC127023 Pan trogl
737	19	67.9	178867	12	AC169205	AC169205 Bos tauru	c 810	19	67.9	233447	12	AC101192	AC101192 Homo sapi
738	19	67.9	179564	6	AC140385	AC140385 Mus muscu	c 811	19	67.9	234714	6	AC110176	AC110176 Mus muscu
739	19	67.9	179564	6	AC140385	AC140385 Mus muscu	c 812	19	67.9	234714	6	AC110176	AC110176 Mus muscu
740	19	67.9	179714	5	AC073941	AC073941 Homo sapi	c 813	19	67.9	235731	12	AC091498	AC091498 Homo sapi
741	19	67.9	179869	12	AC087756	AC087756 Homo sapi	c 814	19	67.9	235839	6	CT009748	CT009748 Mouse DNA
742	19	67.9	180051	6	AC139223	AC139223 Mouse DNA	c 815	19	67.9	235839	11	BX324188	BX324188 Zebrafish
743	19	67.9	181503	5	AL662783	AL662783 Mouse DNA	c 816	19	67.9	235839	11	BX324188	BX324188 Zebrafish
744	19	67.9	182144	5	AC009223	AC009223 Homo sapi	c 817	19	67.9	238312	12	AC114147	AC114147 Rattus no
745	19	67.9	183027	6	AC154458	AC154458 Mus muscu	c 818	19	67.9	238312	12	AC099456	AC099456 Rattus no
746	19	67.9	183238	12	AC158266	AC158266 Orythorh	c 819	19	67.9	239303	12	AC172974	AC172974 Bos tauru
747	19	67.9	183680	5	AC098859	AC098859 Homo sapi	c 820	19	67.9	240836	12	AC108526	AC108526 Rattus no
748	19	67.9	183986	5	AC090660	AC090660 Homo sapi	c 821	19	67.9	241445	12	AC103002	AC103002 Rattus no

822	19	67.9	243489	12	AC096036	895	18.8	67.1	158004	5	AC087319	AC087319 Homo sapi
C 823	19	67.9	244971	6	AL662891	C 896	18.8	67.1	161873	4	AC090056	AC090056 Oryza sat
C 824	19	67.9	245476	12	AC098496	897	18.8	67.1	161938	4	AB006578	AB006578 Oryza sat
C 825	19	67.9	245489	12	AC170939	898	18.8	67.1	166235	6	AC138292	AC138292 Mus muscu
C 826	19	67.9	247805	12	AC131604	899	18.8	67.1	166642	5	AC018919	AC018919 Homo sapi
C 827	19	67.9	250014	12	CR938713	C 900	18.8	67.1	167067	11	BX511268	BX511268 Zebrafish
C 828	19	67.9	250701	5	CR974412	901	18.8	67.1	168062	12	AC012857	AC012857 Homo sapi
C 829	19	67.9	250701	12	CR974436	902	18.8	67.1	168605	12	AC022477	AC022477 Homo sapi
C 830	19	67.9	250804	12	AC097842	C 903	18.8	67.1	168979	12	AC019172	AC019172 Mus muscu
C 831	19	67.9	250906	12	AC123491	C 904	18.8	67.1	168979	11	EX005444	EX005444 Zebrafish
C 832	19	67.9	252271	12	AC157192	C 905	18.8	67.1	169029	5	CNS06C7R	AL390800 Human chr
C 833	19	67.9	257078	12	AC118300	906	18.8	67.1	169884	11	CR848728	CR848728 Zebrafish
C 834	19	67.9	261065	12	AC128913	C 907	18.8	67.1	170813	12	AC181012	AC181012 Strongylo
C 835	19	67.9	263874	12	AC167337	908	18.8	67.1	170908	5	HSBA28009	AL121582 Human DNA
C 836	19	67.9	267058	12	AC134185	C 909	18.8	67.1	172479	13	AC006590	AC006590 Drosophi
C 837	19	67.9	270080	12	AC113644	C 910	18.8	67.1	173688	4	AP005705	AP005705 Oryza sat
C 838	19	67.9	275507	12	AC099381	C 911	18.8	67.1	177660	11	BX276188	BX276188 Zebrafish
C 839	19	67.9	281092	12	AC098047	C 912	18.8	67.1	177770	5	AC110027	AC110027 Homo sapi
C 840	19	67.9	282066	12	AC166668	C 913	18.8	67.1	180916	6	AL844195	AL844195 Mouse DNA
C 841	19	67.9	285689	12	AC160355	914	18.8	67.1	181413	5	AC080032	AC080032 Homo sapi
C 842	19	67.9	287635	12	AC174454	C 915	18.8	67.1	187783	12	AC137462	AC137462 Rattus no
C 843	19	67.9	289818	12	AC108322	C 916	18.8	67.1	188482	12	AC144467	AC144467 Papio anu
C 844	19	67.9	313631	12	AC129821	C 917	18.8	67.1	188876	12	AC162312	AC162312 Mus muscu
C 845	19	67.9	319109	12	AC091368	918	18.8	67.1	189355	11	BX649503	BX649503 Zebrafish
C 846	19	67.9	320208	12	AC157391	C 919	18.8	67.1	194086	6	AC134382	AC134382 Mus muscu
C 847	19	67.9	340000	5	HS21C004	920	18.8	67.1	196129	12	AC167617	AC167617 Bos tauru
C 848	19	67.9	347145	12	AC173518	921	18.8	67.1	196408	12	AC130420	AC130420 Homo sapi
C 849	19	67.9	348540	12	AC132503	922	18.8	67.1	196661	6	AC119967	AC119967 Mus muscu
C 850	19	67.9	349402	15	BX842644	923	18.8	67.1	196758	12	AC009653	AC009653 Homo sapi
C 851	19	67.9	349980	2	AX344566	924	18.8	67.1	199447	11	BX510918	BX510918 Zebrafish
C 852	18.8	67.1	572	13	AY254855S2	925	18.8	67.1	199612	5	AC021582	AC021582 Homo sapi
C 853	18.8	67.1	576	15	NMU93902	C 926	18.8	67.1	201310	6	AC109222	AC109222 Mus muscu
C 854	18.8	67.1	618	15	NMU93903	927	18.8	67.1	204653	6	AC005302	AC005302 Mus muscu
C 855	18.8	67.1	2014	11	BC059513	C 928	18.8	67.1	204720	12	AC008052	AC008052 Homo sapi
C 856	18.8	67.1	2387	13	AY817736	929	18.8	67.1	208170	12	BX901911	BX901911 Danio rer
C 857	18.8	67.1	3780	2	CQ806947	930	18.8	67.1	210130	6	CT009723	CT009723 Mouse DNA
C 858	18.8	67.1	3780	2	CQ807221	931	18.8	67.1	210145	12	AC114314	AC114314 Homo sapi
C 859	18.8	67.1	17758	13	U56959	932	18.8	67.1	210950	6	AC129317	AC129317 Mus muscu
C 860	18.8	67.1	29756	13	U70858	933	18.8	67.1	211817	11	BX649401	BX649401 Zebrafish
C 861	18.8	67.1	55858	6	AL929080	934	18.8	67.1	215455	12	AC123224	AC123224 Rattus no
C 862	18.8	67.1	53727	12	AB023039	C 935	18.8	67.1	220008	6	AC165235	AC165235 Mus muscu
C 863	18.8	67.1	71127	12	AC131278	936	18.8	67.1	223812	6	AC131716	AC131716 Mus muscu
C 864	18.8	67.1	74671	5	AL449106	C 937	18.8	67.1	224213	6	AC147051	AC147051 Homo sapi
C 865	18.8	67.1	75485	11	CR381654	938	18.8	67.1	227600	12	CR956432	CR956432 Mus muscu
C 866	18.8	67.1	87489	12	AC166745	C 939	18.8	67.1	228177	11	BX547928	BX547928 Zebrafish
C 867	18.8	67.1	94773	12	AC14930	940	18.8	67.1	234740	12	AC158693	AC158693 Bos tauru
C 868	18.8	67.1	110000	4	BD061520_3	C 941	18.8	67.1	240675	12	AC106325	AC106325 Rattus no
C 869	18.8	67.1	110000	2	AR409405_3	C 942	18.8	67.1	251681	12	CR788226	CR788226 Danio rer
C 870	18.8	67.1	110000	4	AP008215_063	C 943	18.8	67.1	252440	11	BX005425	BX005425 Zebrafish
C 871	18.8	67.1	110000	4	AP008216_096	C 944	18.8	67.1	256833	12	AC109530	AC109530 Rattus no
C 872	18.8	67.1	110000	4	AE016859_096	C 945	18.8	67.1	260027	13	AE003659	AE003659 Drosophi
C 873	18.8	67.1	110000	12	CR848710_0	946	18.8	67.1	267375	12	AC095339	AC095339 Rattus no
C 874	18.8	67.1	110000	12	TANN4_12	947	18.8	67.1	285218	12	AC112982	AC112982 Bos tauru
C 875	18.8	67.1	110000	15	AE014133_18	948	18.8	67.1	295117	12	AC118895	AC118895 Rattus no
C 876	18.8	67.1	110000	15	BA000003_3	949	18.8	67.1	305666	12	AC166698	AC166698 Bos tauru
C 877	18.8	67.1	114231	5	HS390C10	950	18.8	67.1	330580	12	AC101959	AC101959 Mus muscu
C 878	18.8	67.1	115811	5	AL161787	951	18.8	67.1	348885	12	AC151274	AC151274 Mus muscu
C 879	18.8	67.1	120709	12	AC034185	C 952	18.8	66.4	684	7	BV051539	BV051539 S21P6792
C 880	18.8	67.1	123159	4	AC135798	953	18.6	66.4	684	15	AF062845	AF062845 Escherich
C 881	18.8	67.1	136162	6	AL645994	C 954	18.6	66.4	711	2	AR506781	AR506781 Sequence
C 882	18.8	67.1	138335	6	AC119148	C 955	18.6	66.4	715	7	AR507851	AR507851 Sequence
C 883	18.8	67.1	141705	11	BX571715	956	18.6	66.4	725	7	BV477275	BV477275 G591P6421
C 884	18.8	67.1	143299	5	AL353519	C 957	18.6	66.4	759	3	AM040168	AM040168 Drosophi
C 885	18.8	67.1	144723	5	AC008818	C 958	18.6	66.4	1222	11	AY144593	AY144593 Delima mol
C 886	18.8	67.1	144951	12	AC171445	C 959	18.6	66.4	1495	11	BC072723	BC072723 Danio rer
C 887	18.8	67.1	148468	12	AC160591	C 960	18.6	66.4	1524	4	AF483211	AF483211 Alonosa m
C 888	18.8	67.1	148673	6	AC123048	C 961	18.6	66.4	1574	11	BC095064	BC095064 Danio rer
C 889	18.8	67.1	151365	6	AC127304	C 962	18.6	66.4	1600	11	AY583695	AY583695 Triticum a
C 890	18.8	67.1	153307	6	AL845441	C 963	18.6	66.4	2402	2	CO574074	CO574074 Sequence
C 891	18.8	67.1	154730	13	AC008184	C 964	18.6	66.4	2543	3	AY069621	AY069621 Drosophi
C 892	18.8	67.1	155637	6	AC111012	C 965	18.6	66.4	2624	13	BT033715	BT033715 Drosophi
C 893	18.8	67.1	155992	12	AC083854	C 966	18.6	66.4	2630	13	AY119523	AY119523 Drosophi
C 894	18.8	67.1	155992	12	AC034184	C 967	18.6	66.4	3090	2	AR528057	AR528057 Sequence

ACCESSION AR349062 GI:33749767
VERSION AR349062.1
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 28)
AUTHORS Carter,N.M.
TITLE Nucleic acid probes complementary to human papilloma virus nucleic acid
JOURNAL Patent: US 6583278-A 2 24-JUN-2003;
Gen-Probe Incorporated; San Diego, CA
FEATURES
source location/Qualifiers
1..28
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 100.0%; Score 28; DB 2; Length 28;
Best Local Similarity 100.0%; Pred. No. 0.057;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACATTATGTTATAGTTGTATGGAAC 28
|||||
1 GACATTATGTTATAGTTGTATGGAAC 28

Db 1 GACATTATGTTATAGTTGTATGGAAC 28

RESULT 2
AR349062/c 28 bp DNA linear PAT 17-AUG-2003
LOCUS AR349062
DEFINITION Sequence 2 from patent US 6583278.

ALIGNMENTS

RESULT 1
AR349061 28 bp DNA linear PAT 17-AUG-2003
LOCUS AR349061
DEFINITION Sequence 1 from patent US 6583278.
ACCESSION AR349061
VERSION AR349061.1 GI:33749766
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 28)
AUTHORS Carter,N.M.
TITLE Nucleic acid probes complementary to human papilloma virus nucleic acid
JOURNAL Patent: US 6583278-A 1 24-JUN-2003;
Gen-Probe Incorporated; San Diego, CA
FEATURES
source location/Qualifiers
1..28
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 100.0%; Score 28; DB 2; Length 28;
Best Local Similarity 100.0%; Pred. No. 0.057;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACATTATGTTATAGTTGTATGGAAC 28
|||||
1 GACATTATGTTATAGTTGTATGGAAC 28

Db 1 GACATTATGTTATAGTTGTATGGAAC 28

RESULT 2
AR349062 28 bp DNA linear PAT 17-AUG-2003
LOCUS AR349062
DEFINITION Sequence 2 from patent US 6583278.

ACCESSION AR349062 GI:33749767
VERSION AR349062.1
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 28)
AUTHORS Carter,N.M.
TITLE Nucleic acid probes complementary to human papilloma virus nucleic acid
JOURNAL Patent: US 6583278-A 2 24-JUN-2003;
Gen-Probe Incorporated; San Diego, CA
FEATURES
source location/Qualifiers
1..28
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 100.0%; Score 28; DB 2; Length 28;
Best Local Similarity 100.0%; Pred. No. 0.057;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACATTATGTTATAGTTGTATGGAAC 28
|||||
28 GACATTATGTTATAGTTGTATGGAAC 1

Db 28 GACATTATGTTATAGTTGTATGGAAC 1

RESULT 3
AR349063 28 bp DNA linear PAT 17-AUG-2003
LOCUS AR349063
DEFINITION Sequence 3 from patent US 6583278.
ACCESSION AR349063
VERSION AR349063.1 GI:33749768
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 28)
AUTHORS Carter,N.M.
TITLE Nucleic acid probes complementary to human papilloma virus nucleic acid
JOURNAL Patent: US 6583278-A 3 24-JUN-2003;
Gen-Probe Incorporated; San Diego, CA
FEATURES
source location/Qualifiers
1..28
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 100.0%; Score 28; DB 2; Length 28;
Best Local Similarity 100.0%; Pred. No. 0.057;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACATTATGTTATAGTTGTATGGAAC 28
|||||
28 GACATTATGTTATAGTTGTATGGAAC 1

Db 28 GACATTATGTTATAGTTGTATGGAAC 1

RESULT 4
AR349064 28 bp DNA linear PAT 17-AUG-2003
LOCUS AR349064
DEFINITION Sequence 4 from patent US 6583278.
ACCESSION AR349064
VERSION AR349064.1 GI:33749769
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 28)
AUTHORS Carter,N.M.
TITLE Nucleic acid probes complementary to human papilloma virus nucleic acid
JOURNAL Patent: US 6583278-A 4 24-JUN-2003;

ACCESSION AR349062 GI:33749767
VERSION AR349062.1
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 28)
AUTHORS Carter,N.M.
TITLE Nucleic acid probes complementary to human papilloma virus nucleic acid
JOURNAL Patent: US 6583278-A 2 24-JUN-2003;
Gen-Probe Incorporated; San Diego, CA
FEATURES
source location/Qualifiers
1..28
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 100.0%; Score 28; DB 2; Length 28;
Best Local Similarity 100.0%; Pred. No. 0.057;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACATTATGTTATAGTTGTATGGAAC 28
|||||
1 GACATTATGTTATAGTTGTATGGAAC 28

Db 1 GACATTATGTTATAGTTGTATGGAAC 28

RESULT 2
AR349062/c 28 bp DNA linear PAT 17-AUG-2003
LOCUS AR349062
DEFINITION Sequence 2 from patent US 6583278.

```

FEATURES
  source
    Gen-Probe Incorporated; San Diego, CA
    Location/Qualifiers
      1..28
        /organism="unknown"
        /mol_type="genomic DNA"
ORIGIN
  Query Match
    100.0%; Score 28; DB 2; Length 28;
  Best Local Similarity 100.0%; Pred. No. 0.057;
  Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy
  1 GACATTATTGTTATAGTTGTATGGAAC 28
  |||
  28 GACATTATTGTTATAGTTGTATGGAAC 1

RESULT 5
AR577462 87 bp DNA linear PAT 14-DEC-2004
LOCUS AR577462
DEFINITION Sequence 7 from patent US 6783763.
ACCESSION AR577462
VERSION AR577462.1 GI:56580067
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
  1 (bases 1 to 87)
  Choplin,J., Villada,I.B., Guillet,J.-G., Connan,F. and Ferries,E.
  Polypeptidic proleptic fragments of the E6 and E7 HPV proteins,
  production and use thereof in vaccines
  Patent: US 6783763-A 7 31-AUG-2004;
  Peptide Immune Ligands and Institut National de la Sante et de la
  Recherche Medicale (INSERM); Labège, Cedex;
  FR;
FEATURES
  source
    Location/Qualifiers
      1..87
        /organism="unknown"
        /mol_type="genomic DNA"
ORIGIN
  Query Match
    100.0%; Score 28; DB 2; Length 87;
  Best Local Similarity 100.0%; Pred. No. 0.045;
  Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy
  1 GACATTATTGTTATAGTTGTATGGAAC 28
  |||
  14 GACATTATTGTTATAGTTGTATGGAAC 41

RESULT 6
AX057044 87 bp DNA linear PAT 17-JAN-2001
LOCUS AX057044
DEFINITION Sequence 7 from Patent WO0075336.
ACCESSION AX057044
VERSION AX057044.1 GI:12309885
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
REFERENCE
  1
  Choplin,J., bourgault Villada,I., Guillet,J.-G., Connan,F. and
  Ferries,E.
  Proteifragments die verschiedene epitope der hpv proteine e6 und
  e7 umfassen, ihre herstellung und verwendungen insbesondere zur
  impfung
  Patent: WO 0075336-A 7 14-DEC-2000;
  Biorecator Therapeutics S.A. (FR) ; Institut National de la Sante et
  de la Recherche Medicale (INSERM) (FR)
  Location/Qualifiers
    1..87
      /organism="synthetic construct"
      /mol_type="unassigned DNA"
      /db_xref="taxon:32630"
FEATURES
  source

FEATURES
  source
    /note="fragment de la sequence codant pour E6 de HPV et
    sequence peptidique correspondante"
    1..>87
      /note="unnamed protein product"
      /codon_start=1
      /transl_table=11
      /protein_id="CAC22413.1"
      /db_xref="GI:12309886"
      /translation="MSEYRHYCYSLYGTTLLEQYNNKPLCDLLI"
ORIGIN
  Query Match
    100.0%; Score 28; DB 2; Length 87;
  Best Local Similarity 100.0%; Pred. No. 0.045;
  Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy
  1 GACATTATTGTTATAGTTGTATGGAAC 28
  |||
  14 GACATTATTGTTATAGTTGTATGGAAC 41

RESULT 7
HP014512 271 bp DNA linear VRL 14-SHP-1995
LOCUS HP014512
DEFINITION Human papillomavirus clone HPV1603 E6 protein (E6) gene, partial
  cds.
ACCESSION U14512
VERSION U14512.1 GI:984945
KEYWORDS
SOURCE Human papillomavirus
ORGANISM Human papillomavirus
  Viruses; deDNA viruses, no RNA stage; Papillomaviridae;
  unclassified Papillomaviridae.
  1 (bases 1 to 271)
  Haegert,D.G., Galutira,D.F. and Youngnuband,B.H.
  Sequence variation in the E6 gene of human papillomavirus type 16
  Unpublished
  2 (bases 1 to 271)
  Galutira,D.F.
  Direct Submission
  Submitted (07-SEP-1994) Dante F. Galutira, Memorial University of
  Newfoundland, Discipline of Pathology, Fac. Medicine, St. John's,
  Newfoundland, A1B 3Y6, Canada
FEATURES
  source
    Location/Qualifiers
      1..271
        /organism="Human papillomavirus"
        /mol_type="genomic DNA"
        /strain="HPV16"
        /db_xref="taxon:10566"
        /clone="HPV1603"
        /tissue_type="cervical tissue from invasive carcinoma"
        1..271
          /gene="E6"
          <1..>271
            /gene="E6"
            /codon_start=2
            /product="E6 protein"
            /protein_id="AAB60566.1"
            /db_xref="GI:984946"
            /translation="VYDFAPRDLCTIVRDGNPYAVCDKCLKYSKISRYHYCYSLYG
            TTLEQYNNKPLCDLLIRINCQKPLCPBEKQRHLDKQRPNNRGR"
            85
              /gene="E6"
              /note="T to G transversion from PAL6, GenBank Accession
              Number K02718, position 310; NCBI gi: 333031"
              /replac="c"
ORIGIN
  Query Match
    100.0%; Score 28; DB 10; Length 271;
  Best Local Similarity 100.0%; Pred. No. 0.035;
  Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy
  1 GACATTATTGTTATAGTTGTATGGAAC 28
  |||
  |||
```

Db 108 GACATTATTGTTATGTTGATGGAAC 135

RESULT 8
HPV14513 272 bp DNA linear VRL 15-FEB-2001
LOCUS Human papillomavirus clone HPV1607 E6 protein (E6) gene, partial
DEFINITION cds.
ACCESSION U14513
VERSION U14513.1 GI:984947
KEYWORDS Human papillomavirus
SOURCE Human papillomavirus
ORGANISM Viruses; dsDNA viruses, no RNA stage; Papillomaviridae; unclassified Papillomaviridae.
REFERENCE 1 (bases 1 to 272)
AUTHORS Haegert,D.G., Galutira,D.F. and Youngusband,B.H.
TITLE Sequence variation in the E6 gene of human papillomavirus type 16
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 272)
AUTHORS Galutira,D.F.
TITLE Direct Submission
JOURNAL Submitted (07-SEP-1994) Dante F. Galutira, Memorial University of Newfoundland, Discipline of Pathology, Fac. Medicine, St. John's, Newfoundland, A1B 3V6, Canada
FEATURES
source Location/Qualifiers
1..272
/organism="Human papillomavirus"
/mol_type="genomic DNA"
/strain="HPV16"
/db_xref="taxon:10566"
/clone="HPV1607"
/tissue_type="cervical tissue from invasive carcinoma"
1..272
/gene="E6"
/gene="E6"
/codon_start=2
/product="E6 protein"
/protein_id="AAB60567.2"
/db_xref="GI:12831648"
/translation="VYDFAPRDLWIVRDGNPVAVCDCIKLFYSKISRYHYCSLYGTTLEQYNKPLCDLIRINCQKPLCEBEKQRLDKQRFNIRGR"
variation
31
/gene="E6"
/note="C to G transversion from PA16, GenBank Accession Number K02718, position 256; NCBI gi: 335031"
/replace="C"
ORIGIN
Query Match 100.0%; Score 28; DB 10; Length 272;
Best Local Similarity 100.0%; Pred. No. 0.035;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GACATTATTGTTATGTTGATGGAAC 28
Db 108 GACATTATTGTTATGTTGATGGAAC 135

RESULT 9
AP404692 370 bp DNA linear VRL 09-OCT-2003
LOCUS Human papillomavirus type 16 isolate HPV16BCC1 E6 protein (E6) and E7 protein (E7) genes, partial cds.
DEFINITION AP404692
ACCESSION AP404692
VERSION AP404692.1 GI:15529569
KEYWORDS Human papillomavirus type 16
SOURCE Human papillomavirus type 16
ORGANISM Viruses; dsDNA viruses, no RNA stage; Papillomaviridae; Alphapapillomavirus.
REFERENCE 1 (bases 1 to 370)
AUTHORS Watts,K.J., Thompson,C.H., Coatsart,Y.E. and Rose,B.R.

TITLE Sequence variation and physical state of human papillomavirus type 16 cervical cancer isolates from Australia and New Caledonia
JOURNAL Int. J. Cancer 97 (6), 868-874 (2002)
PUBMED 11857370
REFERENCE 2 (bases 1 to 370)
AUTHORS Watts,K.J., Thompson,C.H., Coatsart,Y.E. and Rose,B.R.
TITLE Direct Submission
JOURNAL Submitted (31-JUL-2001) Department of Infectious Diseases, University of Sydney, Blackburn Building, D06, Off Western Avenue, Camperdown, Sydney, New South Wales 2006, Australia
FEATURES
source Location/Qualifiers
1..370
/organism="Human papillomavirus type 16"
/mol_type="genomic DNA"
/isolate="HPV16BCC1"
/db_xref="taxon:333760"
1..313
/gene="E6"
/gene="E6"
/codon_start=2
/product="E6 protein"
/protein_id="AA01342.1"
/db_xref="GI:15529571"
/translation="DLICIVRDGNPVAVCDCIKLFYSKISRYHYCSLYGTTLEQF NKPLCDLIRINCQKPLCEBEKQRLDKQRFNIRGRWGRGMSCCRSRTREQL"
1..370
/gene="E7"
/gene="E7"
/codon_start=1
/product="E7 protein"
/protein_id="AA01341.1"
/db_xref="GI:15529570"
/translation="WRGDTPLHRYMDLQPE"
ORIGIN
Query Match 100.0%; Score 28; DB 10; Length 370;
Best Local Similarity 100.0%; Pred. No. 0.033;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GACATTATTGTTATGTTGATGGAAC 28
Db 87 GACATTATTGTTATGTTGATGGAAC 114

RESULT 10
S51110 421 bp DNA linear VRL 08-MAY-1993
LOCUS orf E6 [human papillomavirus HPV, type 16, head and neck tumor, Genomic, 421 nt].
DEFINITION S51110
ACCESSION S51110.1 GI:262061
VERSION S51110.1 GI:262061
KEYWORDS Human papillomavirus
SOURCE Human papillomavirus
ORGANISM Viruses; dsDNA viruses, no RNA stage; Papillomaviridae; unclassified Papillomaviridae.
REFERENCE 1 (bases 1 to 421)
AUTHORS Tyan,Y.S., Liu,S.T., Ong,W.R., Chen,M.L., Shu,C.H. and Chang,Y.S.
TITLE Detection of Epstein-Barr virus and human papillomavirus in head and neck tumors
JOURNAL J. Clin. Microbiol. 31 (1), 53-56 (1993)
PUBMED 8380183
REMARK Genbank staff at the National Library of Medicine created this entry [NCBI g1bseq 120826] from the original journal article.
COMMENT PCR primer sequences: 1-21 and 400-421.
FEATURES
source Location/Qualifiers
1..421
/organism="Human papillomavirus"
/mol_type="genomic DNA"
/db_xref="taxon:10566"

gene /note="type: 16"
1..421
/gene="orf E6"
ORIGIN

Query Match 100.0%; Score 28; DB 10; Length 421;
Best Local Similarity 100.0%; Pred. No. 0.032;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACATTATTTGTTATAGTTGTATGGAAC 28
|||||
Db 231 GACATTATTTGTTATAGTTGTATGGAAC 258
|||||

RESULT 11
AF404695 451 bp DNA linear VRL 09-OCT-2003
LOCUS Human papillomavirus type 16 isolate HPV16E6CC4 E6 protein (E6) and
DEFINITION E7 protein (E7) genes, partial cds.
ACCESSION AF404695
VERSION AF404695.1 GI:15529578
KEYWORDS
SOURCE Human papillomavirus type 16
ORGANISM Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
REFERENCE 1 (bases 1 to 451)
AUTHORS Watts,K.J., Thompson,C.H., Cossart,Y.E. and Rose,B.R.
TITLE Sequence variation and physical state of human papillomavirus type 16 cervical cancer isolates from Australia and New Caledonia
JOURNAL Int. J. Cancer 97 (6), 868-874 (2002)
PUBMED 11857370
2 (bases 1 to 451)
AUTHORS Watts,K.J., Thompson,C.H., Cossart,Y.E. and Rose,B.R.
TITLE Direct Submission
JOURNAL Submitted (31-JUL-2001) Department of Infectious Diseases,
University of Sydney, Blackburn Building, D06, Off Western Avenue,
Camperdown, Sydney, New South Wales 2006, Australia
FEATURES
source Location/Qualifiers
1..451
/organism="Human papillomavirus type 16"
/mol_type="genomic DNA"
/isolate="HPV16E6CC4"
/db_xref="taxon:333760"
<1..395
/gene="E6"
<1..395
/gene="E6"
/codon_start=3
/product="E6 protein"
/protein_id="AA01347.1"
/db_xref="GI:15529579"
/translation="TTHDILIECVYCKQQLRREYDFAFRDLCTVYRDGNPYAVCDK
CLIKYYSKISEYRHYCYSLYGTLLIEQYKPKLCDLLIRICINCKPLCPBEKORHLDKQ
RFHNI RGRWTRCSCRSRRTRETQ."
398..>451
/gene="E7"
398..>451
/gene="E7"
/codon_start=1
/product="E7 protein"
/protein_id="AA01348.1"
/db_xref="GI:15529580"
/translation="WHGDTPTLHMYMDLQPE"

ORIGIN

Query Match 100.0%; Score 28; DB 10; Length 451;
Best Local Similarity 100.0%; Pred. No. 0.032;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACATTATTTGTTATAGTTGTATGGAAC 28
|||||
Db 169 GACATTATTTGTTATAGTTGTATGGAAC 196
|||||

RESULT 12
AF404703 451 bp DNA linear VRL 09-OCT-2003
LOCUS Human papillomavirus type 16 isolate HPV16E6CC12 E6 protein (E6)
DEFINITION and E7 protein (E7) genes, partial cds.
ACCESSION AF404703
VERSION AF404703.1 GI:15529602
KEYWORDS
SOURCE Human papillomavirus type 16
ORGANISM Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
REFERENCE 1 (bases 1 to 451)
AUTHORS Watts,K.J., Thompson,C.H., Cossart,Y.E. and Rose,B.R.
TITLE Sequence variation and physical state of human papillomavirus type 16 cervical cancer isolates from Australia and New Caledonia
JOURNAL Int. J. Cancer 97 (6), 868-874 (2002)
PUBMED 11857370
2 (bases 1 to 451)
AUTHORS Watts,K.J., Thompson,C.H., Cossart,Y.E. and Rose,B.R.
TITLE Direct Submission
JOURNAL Submitted (31-JUL-2001) Department of Infectious Diseases,
University of Sydney, Blackburn Building, D06, Off Western Avenue,
Camperdown, Sydney, New South Wales 2006, Australia
FEATURES
source Location/Qualifiers
1..451
/organism="Human papillomavirus type 16"
/mol_type="genomic DNA"
/isolate="HPV16E6CC12"
/db_xref="taxon:333760"
<1..395
/gene="E6"
<1..395
/gene="E6"
/codon_start=3
/product="E6 protein"
/protein_id="AA01363.1"
/db_xref="GI:15529603"
/translation="TTHNIILIECVYCKQQLRREYDFAFRDLCTVYRDGNPYAVCDK
CLIKYYSKISEYRHYCYSLYGTLLIEQYKPKLCDLLIRICINCKPLCPBEKORHLDKQ
RFHNI RGRWTRCSCRSRRTRETQ."
398..>451
/gene="E7"
398..>451
/gene="E7"
/codon_start=1
/product="E7 protein"
/protein_id="AA01364.1"
/db_xref="GI:15529604"
/translation="WHGDTPTLHMYMDLQPE"

ORIGIN

Query Match 100.0%; Score 28; DB 10; Length 451;
Best Local Similarity 100.0%; Pred. No. 0.032;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACATTATTTGTTATAGTTGTATGGAAC 28
|||||
Db 169 GACATTATTTGTTATAGTTGTATGGAAC 196
|||||

RESULT 13
HPU34111 456 bp DNA linear VRL 08-MAR-1996
LOCUS Human papillomavirus type 16, isolate NM T455, early transforming
DEFINITION protein E6 (E6) gene, complete cds.
ACCESSION U34111
VERSION U34111.1 GI:1098727
KEYWORDS
SOURCE Human papillomavirus
ORGANISM Human papillomavirus

REFERENCE	1	Viruses: dsDNA viruses, no RNA stage; Papillomaviridae.
AUTHORS	Yamada,T., Wheeler,C.M., Halpern,A.L., Stewart,A.C., Hildesheim,A. and Jenison,S.A.	
TITLE	Human papillomavirus type 16 variant lineages in United States populations characterized by nucleotide sequence analysis of the E6, E2, and L1 coding segments	
JOURNAL	J. Virol. 69 (12), 7743-7753 (1995)	
PUBMED	7494284	
REFERENCE	2 (bases 1 to 456)	
AUTHORS	Farmer,A.D.	
TITLE	Direct Submission	
JOURNAL	Submitted (16-AUG-1995) Andrew D. Farmer, Los Alamos National Laboratory, HIV Sequence Database, T-10, Mail Stop K710, Los Alamos, NM 87501, USA	
FEATURES	Location/Qualifiers	
source	1..456	
organism	"Human papillomavirus"	
molecule	"mol type="Genomic DNA"	
isolate	"isolate="NM T455"	
db_xref	"db_xref="taxon:10566"	
note	"note="Collected by cervical swab at the University of New Mexico, New Mexico, United States."	
gene	1..456	
gene	"gene="E6"	
cds	1..456	
cds	"/gene="E6"	
cds	"/codon_start=1	
cds	"/product="early transforming protein E6"	
cds	"/protein_id="AA91658.1"	
cds	"/db_xref="GI:1098728"	
cds	"/translation="MFQDPQPRPKLQPOLCTELQTHIDTIIIELCVYKQQLRREYDFAFDLCIVYDGNPVACDCKLFYKISYRVCYSLGTLEQQYNPLDILRCINQKPLCPERKQRLDKQRFNIRKGRMTGRMCSGCRSSRTREYQL"	
ORIGIN		
Query Match	100.0%; Score 28; DB 10; Length 456;	
Best Local Similarity	100.0%; Pred. No. 0.032;	
Matches	28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 GACATTATTGTTATGTTGATGAGAC 28	
DB	230 GACATTATTGTTATGTTGATGAGAC 257	
RESULT 14		
HPV34115	456 bp DNA linear VRL 08-MAR-1996	
LOCUS	Human papillomavirus type 16, isolate OR 2087, early transforming	
DEFINITION	Protein E6 (E6) gene, complete cde.	
ACCESSION	U34115	
VERSION	U34115.1 GI:1098735	
KEYWORDS		
SOURCE	Human papillomavirus	
ORGANISM	Human papillomavirus	
REFERENCE	1 (bases 1 to 456)	
AUTHORS	Yamada,T., Wheeler,C.M., Halpern,A.L., Stewart,A.C., Hildesheim,A. and Jenison,S.A.	
TITLE	Human papillomavirus type 16 variant lineages in United States populations characterized by nucleotide sequence analysis of the E6, E2, and L1 coding segments	
JOURNAL	J. Virol. 69 (12), 7743-7753 (1995)	
PUBMED	7494284	
REFERENCE	2 (bases 1 to 456)	
AUTHORS	Farmer,A.D.	
TITLE	Direct Submission	
JOURNAL	Submitted (16-AUG-1995) Andrew D. Farmer, Los Alamos National Laboratory, HIV Sequence Database, T-10, Mail Stop K710, Los Alamos, NM 87501, USA	
FEATURES	Location/Qualifiers	

FEATURES	source
gene	<p>1. .456</p> <p>/organism="Human papillomavirus"</p> <p>/mol_type="genomic DNA"</p> <p>/isolate="OR 2087"</p> <p>/db_xref="taxon:10566"</p> <p>/note="collected by cervical lavage at Kaiser Permanente, Portland, Oregon, United States."</p> <p>1. .456</p> <p>/gene="B6"</p> <p>1. .456</p> <p>/gene="B6"</p> <p>/codon_start=1</p> <p>/product="early transforming protein E6"</p> <p>/protein_id="AA91662.1"</p> <p>/db_xref="GI:1098736"</p> <p>/translation="MFQDDPERPRKLPOLCTELQTTIHIIIECVYCKQOQLLRREYDFAKRDICIVYDGNPAVACDKCFKFSKISRYHVCYSLSYGTLLBEOQYKPKCDLLIR</p>
CDS	<p>1. .456</p> <p>/gene="B6"</p> <p>1. .456</p> <p>/gene="B6"</p> <p>/codon_start=1</p> <p>/product="early transforming protein E6"</p> <p>/protein_id="AA91662.1"</p> <p>/db_xref="GI:1098736"</p> <p>/translation="MFQDDPERPRKLPOLCTELQTTIHIIIECVYCKQOQLLRREYDFAKRDICIVYDGNPAVACDKCFKFSKISRYHVCYSLSYGTLLBEOQYKPKCDLLIR</p>
gene	<p>1. .456</p> <p>/organism="Human papillomavirus"</p> <p>/mol_type="genomic DNA"</p> <p>/isolate="OR 2087"</p> <p>/db_xref="taxon:10566"</p> <p>/note="collected by cervical lavage at Kaiser Permanente, Portland, Oregon, United States."</p> <p>1. .456</p> <p>/gene="B6"</p> <p>1. .456</p> <p>/gene="B6"</p> <p>/codon_start=1</p> <p>/product="early transforming protein E6"</p> <p>/protein_id="AA91662.1"</p> <p>/db_xref="GI:1098736"</p> <p>/translation="MFQDDPERPRKLPOLCTELQTTIHIIIECVYCKQOQLLRREYDFAKRDICIVYDGNPAVACDKCFKFSKISRYHVCYSLSYGTLLBEOQYKPKCDLLIR</p>
CDS	<p>1. .456</p> <p>/gene="B6"</p> <p>1. .456</p> <p>/gene="B6"</p> <p>/codon_start=1</p> <p>/product="early transforming protein E6"</p> <p>/protein_id="AA91662.1"</p> <p>/db_xref="GI:1098736"</p> <p>/translation="MFQDDPERPRKLPOLCTELQTTIHIIIECVYCKQOQLLRREYDFAKRDICIVYDGNPAVACDKCFKFSKISRYHVCYSLSYGTLLBEOQYKPKCDLLIR</p>

ORIGIN	CINCOKPLCPREKORHLDKQRFHNIRGRWGRGMSCCRSRTRERQL"
Query Match	100.0%; Score 28; DB 10; Length 456;
Best Local Similarity	100.0%; Pred. No. 0.032;
Matches	28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy	1 GACATTATGTTAGTTGTATGGAAC 28
Db	230 GACATTATGTTAGTTGTATGGAAC 257
RESULT 16	
HPJ34121	
LOCUS	HPJ34121 456 bp DNA linear VRL 08-MAR-1996
DEFINITION	Human papillomavirus type 16, isolate OR 4724, early transforming protein E6 (E6) gene, complete cds.
ACCESSION	U34121
VERSION	U34121.1
KEYWORDS	GI:1098747
SOURCE	Human papillomavirus
ORGANISM	Human papillomavirus
REFERENCE	Yamada,T., Wheeler,C.M., Halpern,A.L., Stewart,A.C., Hildesheim,A. and Jenison,S.A.
AUTHORS	Human papillomavirus type 16 variant lineages in United States populations characterized by nucleotide sequence analysis of the E6, E2, and E1 coding segments
TITLE	J. Virol. 69 (12), 7743-7753 (1995)
JOURNAL	7494284
PUBMED	2 (bases 1 to 456)
REFERENCE	Farmer,A.D.
AUTHORS	Direct Submission
TITLE	Submitted (16-AUG-1995) Andrew D. Farmer, Los Alamos National Laboratory, HIV Sequence Database, T-10, Mail Stop K710, Los Alamos, NM 87501, USA
JOURNAL	Location/Qualifiers
FEATURES	1..456
source	/organism="Human papillomavirus"
	/mol type="genomic DNA"
	/isolate="OR 4724"
	/db_xref="taxon:10566"
	/note="collected by cervical lavage at Kaiser Permanente, Portland, Oregon, United States."
gene	1..456
	/gene="E6"
CDS	1..456
	/gene="E6"
	/codon_start=1
	/product="early transforming protein E6"
	/protein_id="AA091668.1"
	/db_xref="GI:1098748"
	/translation="MFQDPERPKLPLQCTELQTTIHDIILCEVCYKQQLRREYVD FAPFDLCIVRDNPAVCDCIKFYSKISRYHYCYSLYGTLEQDYNPLCDLLIR CINQKPLCPREKORHLDKQRFHNIRGRWGRGMSCCRSRTRERQL"
ORIGIN	
Query Match	100.0%; Score 28; DB 10; Length 456;
Best Local Similarity	100.0%; Pred. No. 0.032;
Matches	28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy	1 GACATTATGTTAGTTGTATGGAAC 28
Db	230 GACATTATGTTAGTTGTATGGAAC 257
RESULT 17	
HPJ34123	
LOCUS	HPJ34123 456 bp DNA linear VRL 08-MAR-1996
DEFINITION	Human papillomavirus type 16, isolate OR 5110, early transforming protein E6 (E6) gene, complete cds.

VERSION	U34123
KEYWORDS	U34123.1 GI:1098751
SOURCE	
ORGANISM	Human papillomavirus Human papillomavirus viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
REFERENCE	1 (bases 1 to 456) Yamada,T., Wheeler,C.M., Halpern,A.L., Stewart,A.C., Hildesheim,A. and Jensen,S.A.
TITLE	Human papillomavirus type 16 variant lineages in United States populations characterized by nucleotide sequence analysis of the E6, E2, and L1 coding segments J Virol. 69 (12), 7743-7753 (1995)
JOURNAL	7494284
PUBMED	2 (bases 1 to 456) Farmer,A.D.
REFERENCE	Direct Submission Submitted (16-AUG-1995) Andrew D. Farmer, Los Alamos National Laboratory, HIV Sequence Database, T-10, Mail Stop K710, Los Alamos, NM 87501, USA
TITLE	Location/Qualifiers
JOURNAL	1. .456 /organism="Human papillomavirus" /mol_type="genomic DNA" /isolate="OR 5110" /db_xref="taxon:10566" /note="Collected by cervical lavage at Kaiser Permanente, Portland, Oregon, United States." 1. .456 /gene="E6" 1. .456 /gene="E6"
CDS	/codon_start=1 /product="early transforming protein E6" /protein_id="AA91670.1" /db_xref="GI:1098752" /translation="MFQDQERPRKLPQLCTELIQTTHIIIECYCKQQGLRRBYD FARDIDIVRDGNPNVADQDKLKFYSKISEYRHYCSLYGTLLEOQYNPLCDLLIR CINCRPFLCEBKQRHLDKKRFHNIRGWTRGCMSCSSRTTRETOL"
ORIGIN	
Query Match	100.0%; Score 28; DB 10; Length 456; Best Local Similarity 100.0%; Pred.No.0.032; Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Cy	1 GACATTATGTATTAGTTTGATGAAC 28
Db	230 GACATTATGTATTAGTTTGATGAAC 257
RESULT 18	
HPU34124	456 bp DNA linear VRL 08-MAR-1996
LOCUS	HPU34124
DEFINITION	Human papillomavirus type 16, isolate OR 5428, early transforming protein E6 (E6) gene, complete cds.
ACCESSION	U34124
VERSION	U34124.1 GI:1098753
KEYWORDS	Human papillomavirus Human papillomavirus viruses; dsDNA viruses, no RNA stage; Papillomaviridae; unclassified Papillomaviridae.
SOURCE	1 (bases 1 to 456) Yamada,T., Wheeler,C.M., Halpern,A.L., Stewart,A.C., Hildesheim,A. and Jensen,S.A.
ORGANISM	Human papillomavirus type 16 variant lineages in United States populations characterized by nucleotide sequence analysis of the E6, E2, and L1 coding segments J Virol. 69 (12), 7743-7753 (1995)
TITLE	7494284
JOURNAL	2 (bases 1 to 456) Farmer,A.D.
PUBMED	
REFERENCE	
AUTHORS	
REFERENCES	
OTHERS	

TITLE Direct Submission
JOURNAL Submitted (16-AUG-1995) Andrew D. Farmer, Los Alamos National Laboratory, HIV Sequence Database, T-10, Mail Stop K710, Los Alamos, NM 87501, USA

FEATURES
source
1..456
/organism="Human papillomavirus"
/mol_type="genomic DNA"
/isolate="OR 5428"
/db_xref="taxon:10566"
/note="collected by cervical lavage at Kaiser Permanente, Portland, Oregon, United States."
1..456
/gene="E6"
1..456
/gene="E6"
/codon_start=1
/product="early transforming protein E6"
/protein_id="AA91671.1"
/db_xref="GI:1098754"
/translation="MFQDQERPRKLPQLCTELQTTIHETIECYCKQQLRRBYDFAFRLCIVRDGNPVAVCDKCLFKYSISRYRHCYSLYGTTLAQNKPLCDLLIR CINCKPLCPBEKQRLDKQRFNIRGRWTRGCMSCCRSRTRETLQ"

ORIGIN
Query Match 100.0%; Score 28; DB 10; Length 456;
Best Local Similarity 100.0%; Pred. No. 0.032;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACATTATGTTATGTTGTATGGAAC 28
|||||
230 GACATTATGTTATGTTGTATGGAAC 257

Db

RESULT 19
HPV34126 456 bp DNA linear VRL 08-MAR-1996
LOCUS Human papillomavirus type 16, isolate OR 6170, early transforming protein E6 (E6) gene, complete cds.
DEFINITION U34126
ACCESSION U34126
VERSION U34126.1 GI:1098757
KEYWORDS
SOURCE Human papillomavirus
ORGANISM Human papillomavirus
VIRUSES; dsDNA viruses, no RNA stage; Papillomaviridae;
unclassified Papillomaviridae.
REFERENCE 1 (bases 1 to 456)
Yamada,T., Wheeler,C.M., Halpern,A.L., Stewart,A.C., Hildesheim,A. and Jensen,S.A.
Human papillomavirus type 16 variant lineages in United States populations characterized by nucleotide sequence analysis of the E6, L2, and L1 coding segments
J. Virol. 69 (12), 7743-7753 (1995)
7494284
2 (bases 1 to 456)
Farmer,A.D.
Direct Submission
Submitted (16-AUG-1995) Andrew D. Farmer, Los Alamos National Laboratory, HIV Sequence Database, T-10, Mail Stop K710, Los Alamos, NM 87501, USA
Location/Qualifiers
1..456
/organism="Human papillomavirus"
/mol_type="genomic DNA"
/isolate="OR 6170"
/db_xref="taxon:10566"
/note="collected by cervical lavage at Kaiser Permanente, Portland, Oregon, United States."
1..456
/gene="E6"
1..456
/gene="E6"
/codon_start=1

gene
CDS

/product="early transforming protein E6"
/protein_id="AA91673.1"
/db_xref="GI:1098758"
/translation="MFQDQERPRKLPQLCTELQTTIHETIECYCKQQLRRBYDFAFRLCIVRDGNPVAVCDKCLFKYSISRYRHCYSLYGTTLAQNKPLCDLLIR CINCKPLCPBEKQRLDKQRFNIRGRWTRGCMSCCRSRTRETLQ"

ORIGIN
Query Match 100.0%; Score 28; DB 10; Length 456;
Best Local Similarity 100.0%; Pred. No. 0.032;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACATTATGTTATGTTGTATGGAAC 28
|||||
230 GACATTATGTTATGTTGTATGGAAC 257

Db

RESULT 20
HPV34129 456 bp DNA linear VRL 08-MAR-1996
LOCUS Human papillomavirus type 16, isolate OR 7574, early transforming protein E6 (E6) gene, complete cds.
DEFINITION U34129
ACCESSION U34129
VERSION U34129.1 GI:1098763
KEYWORDS
SOURCE Human papillomavirus
ORGANISM Human papillomavirus
VIRUSES; dsDNA viruses, no RNA stage; Papillomaviridae;
unclassified Papillomaviridae.
REFERENCE 1 (bases 1 to 456)
Yamada,T., Wheeler,C.M., Halpern,A.L., Stewart,A.C., Hildesheim,A. and Jensen,S.A.
Human papillomavirus type 16 variant lineages in United States populations characterized by nucleotide sequence analysis of the E6, L2, and L1 coding segments
J. Virol. 69 (12), 7743-7753 (1995)
7494284
2 (bases 1 to 456)
Farmer,A.D.
Direct Submission
Submitted (16-AUG-1995) Andrew D. Farmer, Los Alamos National Laboratory, HIV Sequence Database, T-10, Mail Stop K710, Los Alamos, NM 87501, USA
Location/Qualifiers
1..456
/organism="Human papillomavirus"
/mol_type="genomic DNA"
/isolate="OR 7574"
/db_xref="taxon:10566"
/note="collected by cervical lavage at Kaiser Permanente, Portland, Oregon, United States."
1..456
/gene="E6"
1..456
/gene="E6"
/codon_start=1
/product="early transforming protein E6"
/protein_id="AA91676.1"
/db_xref="GI:1098764"
/translation="MFQDQERPRKLPQLCTELQTTIHETIECYCKQQLRRBYDFAFRLCIVRDGNPVAVCDKCLFKYSISRYRHCYSLYGTTLAQNKPLCDLLIR CINCKPLCPBEKQRLDKQRFNIRGRWTRGCMSCCRSRTRETLQ"

ORIGIN
Query Match 100.0%; Score 28; DB 10; Length 456;
Best Local Similarity 100.0%; Pred. No. 0.032;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACATTATGTTATGTTGTATGGAAC 28
|||||
230 GACATTATGTTATGTTGTATGGAAC 257

Db

```

RESULT 21
AF003016          458 bp   DNA      linear   VRL 20-SEP-1997
LOCUS              Human papillomavirus type 16 E6 protein gene, complete cds.
DEFINITION
ACCESSION          AF003016
VERSION            AF003016.1 GI:2196714
KEYWORDS
SOURCE
ORGANISM            Human papillomavirus type 16
                    Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
                    Alphapapillomavirus.
REFERENCE
AUTHORS            Tornesello,M.L., Buonaguro,F.M., Meglio,A., Buonaguro,L.,
                    Beth-Girardo,E. and Girardo,G.
TITLE              Sequence variations and viral genomic state of human papillomavirus
                    type 16 in penile carcinomas from Ugandan patients
JOURNAL            J. Gen. Virol. 78 (Pt 9), 2199-2208 (1997)
PUBMED            9292007
REFERENCE
AUTHORS            Tornesello,M.L., Buonaguro,F.M., Meglio,A., Buonaguro,L.,
                    Beth-Girardo,E. and Girardo,G.
TITLE              Direct Submission
                    Submitted (06-MAY-1997) Viral Oncology, National Cancer Institute,
                    Capella Cangiani, Naples I-80131, Italy
JOURNAL
FEATURES
source              1. 458
                    /organism="Human papillomavirus type 16"
                    /mol_type="genomic DNA"
                    /isolate="PC-7"
                    /db_xref="taxon:333760"
                    /note="Ugandan sample"
                    3..458
                    /codon_start=1
                    /product="E6 protein"
                    /protein_id="AAB70733.1"
                    /db_xref="GI:2196715"
                    /translation="MFQDPOERPTKLPDLCTELQTTIHDIIECVYCKQOQLRREYVD
                    FAFRDLCIVRDGNPVAVCDKLFKYSKISRYHVCYSLYGTLLBOQYKPLCDLILIR
                    CINGQKPLCPDEKQRLDKKQRFNIRGWTGRGCMSCSSRTRRETQL"

ORIGIN
Query Match          100.0%; Score 28; DB 10; Length 458;
Best Local Similarity 100.0%; Pred. No. 0.032;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACATTATTGTTATAGTTGTATGGAAC 28
Db 232 GACATTATTGTTATAGTTGTATGGAAC 259

RESULT 22
AR577459          477 bp   DNA      linear   PAT 14-DEC-2004
LOCUS              AR577459
DEFINITION          Sequence 1 from patent US 6783763.
ACCESSION          AR577459
VERSION            AR577459.1 GI:56580064
KEYWORDS
SOURCE
ORGANISM            Unknown.
                    Unclasseified.
                    1 (bases 1 to 477)
REFERENCE
AUTHORS            Choppin,J., Villada,I.B., Guillet,J.-G., Connan,F. and Ferries,E.
TITLE              Polypeptidic proteic fragments of the E6 and E7 HPV proteins,
                    production and use thereof in vaccines
JOURNAL            Patent: US 6783763-A 1 31-AGG-2004;
                    Peptide Immune Ligands and Institut National de la Sante et de la
                    Recherche Medicale (INSERM); Labège, Cedex;
                    FR;
FEATURES
source              1..477
                    /organism="unknown"
                    /mol_type="genomic DNA"

```

```

Query Match          100.0%; Score 28; DB 2; Length 477;
Best Local Similarity 100.0%; Pred. No. 0.031;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACATTATTGTTATAGTTGTATGGAAC 28
Db 251 GACATTATTGTTATAGTTGTATGGAAC 278

RESULT 23
AX057038          477 bp   DNA      linear   PAT 17-JAN-2001
LOCUS              AX057038
DEFINITION          Sequence 1 from Patent WO0075336.
ACCESSION          AX057038
VERSION            AX057038.1 GI:12309879
KEYWORDS
SOURCE
ORGANISM            Human papillomavirus
                    Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
                    unclassified Papillomaviridae.
REFERENCE
AUTHORS            Choppin,J., Bourgault Villada,I., Guillet,J.-G., Connan,F. and
                    Ferries,E.
TITLE              Proteinfraemente die verschiedene epitope der hpv proteine e6 und
                    e7 umfassen, ihre herstellung und verwendungen insbesondere zur
                    impfung
JOURNAL            Patent: WO 0075336-A 1 14-DEC-2000;
                    Biovector Therapeutics S.A. (FR) ; Institut National de la Sante et
                    de la Recherche Medicale (INSERM) (FR)
FEATURES
source              1. 477
                    /organism="Human papillomavirus"
                    /mol_type="unassigned DNA"
                    /db_xref="taxon:10566"
                    1..477
                    /note="unnamed protein product"
                    /codon_start=1
                    /protein_id="CAC22410.1"
                    /db_xref="GI:12309880"
                    /translation="MHQKRTAMFQDPOERPRKLPQLCTELQTTIHDIIECVYCKQOQL
                    LRREYVDPAFRDLCIVRDGNPVAVCDKLFKYSKISRYHVCYSLYGTLLBOQYKPLCDLILIR
                    CINGQKPLCPDEKQRLDKKQRFNIRGWTGRGCMSCSSRTRRETQL"

ORIGIN
Query Match          100.0%; Score 28; DB 2; Length 477;
Best Local Similarity 100.0%; Pred. No. 0.031;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACATTATTGTTATAGTTGTATGGAAC 28
Db 251 GACATTATTGTTATAGTTGTATGGAAC 278

RESULT 24
AX766063          477 bp   DNA      linear   PAT 25-JUN-2003
LOCUS              AX766063
DEFINITION          Sequence 7 from Patent WO03008573.
ACCESSION          AX766063
VERSION            AX766063.1 GI:32260136
KEYWORDS
SOURCE
ORGANISM            Human papillomavirus type 16
                    Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
                    Alphapapillomavirus.
REFERENCE
AUTHORS            Milner,A.J.
TITLE              Silencing of gene expression
JOURNAL            Patent: WO 03008573-A 7 30-JAN-2003;
                    Milner, Anne Josephine (GB)
FEATURES
source              1..477
                    /organism="Human papillomavirus type 16"

```

```

/mot_type="unassigned DNA"
/db_xref="taxon:333760"
ORIGIN
Query Match      100.0%; Score 28; DB 2; Length 477;
Best Local Similarity 100.0%; Pred. No. 0.031;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY
1 GACATTATGTTAGTTGATGGAAC 28
|||||
251 GACATTATGTTAGTTGATGGAAC 278

RESULT 25
AF486299      477 bp  DNA      linear  VRL 12-AUG-2002
LOCUS      Human papillomavirus type 16 strain As-P isolate PMH-Q24 E6 protein
DEFINITION      (E6) gene, complete cds.
ACCESSION      AF486299
VERSION
KEYWORDS
SOURCE
ORGANISM
Human papillomavirus type 16
Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
Alphapapillomavirus.
1 (bases 1 to 477)
Chan, P.K.S., Lam, C.W., Cheung, T.H., Li, W.W.H., Lo, K.W.K.,
Chan, M.Y.M., Cheung, J.L.K., Xu, L.Y. and Cheng, A.F.
Human papillomavirus type 16 intertypic variant infection and risk
for cervical neoplasia in Southern China
J. Infect. Dis. 186 (5), 696-700 (2002)
2 (bases 1 to 477)
Chan, P.K.S., Lam, C.W., Cheung, T.H., Li, W.W.H., Lo, K.W.K.,
Chan, M.Y.M., Cheung, J.L.K., Xu, L.Y. and Cheng, A.F.
Direct Submision
Submitted (22-FEB-2002) Department of Microbiology, The Chinese
University of Hong Kong, Prince of Wales Hospital, Shatin, N.T.,
Hong Kong SAR, China
FEATURES
source
location/Qualifiers
1..477
/organism="Human papillomavirus type 16"
/mol_type="genomic DNA"
/strain="As-P"
/isolate="PMH-Q24"
/db_xref="taxon:333760"
/country="China"
/note="Isolated from a patient with invasive squamous cell
carcinoma of the cervix"
1..477
/gene="E6"
1..477
/gene="E6"
/note="oncprotein; transforming; trans-activating;
p53-binding"
/codon_start=1
/product="E6 protein"
/protein_id="AAL96604.1"
/db_xref="GI:19744648"
/translation="MHQRTAMFODPQPERPKLPOLCTELQTHIEIILBCYCKQQL
LRREYDPAFRLCIVRDGNPVAVCDKLFYSKISRYRHVCYSLYGTTLQOYNKP
LCDLIRICINQKPLCEPKRKHLDKQRFNINRGRWGRGMSCCRSRTRRTQL"

ORIGIN
Query Match      100.0%; Score 28; DB 10; Length 477;
Best Local Similarity 100.0%; Pred. No. 0.031;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY
1 GACATTATGTTAGTTGATGGAAC 28
|||||
251 GACATTATGTTAGTTGATGGAAC 278

RESULT 26

```

```

AF486300      477 bp  DNA      linear  VRL 12-AUG-2002
LOCUS      Human papillomavirus type 16 strain As-P isolate PMH-Q59 E6 protein
DEFINITION      (E6) gene, complete cds.
ACCESSION      AF486300
VERSION      AF486300.1 GI:19744649
KEYWORDS
SOURCE
ORGANISM
Human papillomavirus type 16
Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
Alphapapillomavirus.
1 (bases 1 to 477)
Chan, P.K.S., Lam, C.W., Cheung, T.H., Li, W.W.H., Lo, K.W.K.,
Chan, M.Y.M., Cheung, J.L.K., Xu, L.Y. and Cheng, A.F.
Human papillomavirus type 16 intertypic variant infection and risk
for cervical neoplasia in Southern China
J. Infect. Dis. 186 (5), 696-700 (2002)
2 (bases 1 to 477)
Chan, P.K.S., Lam, C.W., Cheung, T.H., Li, W.W.H., Lo, K.W.K.,
Chan, M.Y.M., Cheung, J.L.K., Xu, L.Y. and Cheng, A.F.
Direct Submision
Submitted (22-FEB-2002) Department of Microbiology, The Chinese
University of Hong Kong, Prince of Wales Hospital, Shatin, N.T.,
Hong Kong SAR, China
FEATURES
source
location/Qualifiers
1..477
/organism="Human papillomavirus type 16"
/mol_type="genomic DNA"
/strain="As-P"
/isolate="PMH-Q59"
/db_xref="taxon:333760"
/country="China"
/note="Isolated from a patient with invasive squamous cell
carcinoma of the cervix"
1..477
/gene="E6"
1..477
/gene="E6"
/note="oncprotein; transforming; trans-activating;
p53-binding"
/codon_start=1
/product="E6 protein"
/protein_id="AAL96605.1"
/db_xref="GI:19744650"
/translation="MHQRTAMFODPQPERPKLPOLCTELQTHIEIILBCYCKQQL
LRREYDPAFRLCIVRDGNPVAVCDKLFYSKISRYRHVCYSLYGTTLQOYNKP
LCDLIRICINQKPLCEPKRKHLDKQRFNINRGRWGRGMSCCRSRTRRTQL"

ORIGIN
Query Match      100.0%; Score 28; DB 10; Length 477;
Best Local Similarity 100.0%; Pred. No. 0.031;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY
1 GACATTATGTTAGTTGATGGAAC 28
|||||
251 GACATTATGTTAGTTGATGGAAC 278

RESULT 27
AF486301      477 bp  DNA      linear  VRL 12-AUG-2002
LOCUS      Human papillomavirus type 16 strain As-P isolate OER-2076 E6
DEFINITION      (E6) gene, complete cds.
ACCESSION      AF486301
VERSION      AF486301.1 GI:19744651
KEYWORDS
SOURCE
ORGANISM
Human papillomavirus type 16
Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
Alphapapillomavirus.
1 (bases 1 to 477)
Chan, P.K.S., Lam, C.W., Cheung, T.H., Li, W.W.H., Lo, K.W.K.,
Chan, M.Y.M., Cheung, J.L.K., Xu, L.Y. and Cheng, A.F.

```

```

TITLE      Human papillomavirus type 16 intraepithelial variant infection and risk
JOURNAL    J. Infect. Dis. 186 (5), 696-700 (2002)
REFERENCE  2. (bases 1 to 477)
AUTHORS    Chan, P.K.S., Lam, C.W., Cheung, T.H., Li, W.W.H., Lo, K.W.K.,
            Chan, M.Y.M., Cheung, J.L.K., Xu, L.Y. and Cheng, A.F.
TITLE      Direct Submision
JOURNAL    Submitted (22-FEB-2002) Department of Microbiology, The Chinese
            University of Hong Kong, Prince of Wales Hospital, Shatin, N.T.,
            Hong Kong SAR, China

FEATURES   source
            1..477
            /organism="Human papillomavirus type 16"
            /mol_type="genomic DNA"
            /strain="As-P"
            /isolate="QEH-2076"
            /db_xref="taxon:333760"
            /country="China"
            /note="Isolated from a patient with cervical
            intraepithelial neoplasia grade II"

gene       1..477
            /gene="E6"
            /gene="E6"
            1..477
            /note="oncoprotein; transforming; trans-activating;
            p53-binding"
            /product="E6 protein"
            /protein_id="AAL96607.1"
            /db_xref="GI:19744652"
            /translation="MHOKRTAMFODPOERPRKLPOLCTELQTTIHIIIECYCKQOL
            LRREYDPAFLDICTVYRDGNPVAQCDKCLKYSKISSEYRHVCYSLYGTTLQOYKPK
            LCDLIRICNCKPLCPBEKQHLDKQRFNIRGRWGRGCMSCCRSSRTRETL"

ORIGIN
Query Match      100.0%; Score 28; DB 10; Length 477;
Best Local Similarity 100.0%; Pred. No. 0.031;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 GACATTATTGTTATAGTTGTATGGAAC 28
        |||||||
        251 GACATTATTGTTATAGTTGTATGGAAC 278

RESULT 28
AF486302      477 bp      DNA      linear      VRL 12-AUG-2002
LOCUS        Human papillomavirus type 16 strain As-P isolate PWH-Q81 E6 protein
DEFINITION   (E6) gene, complete cds.
ACCESSION    AF486302
VERSION      AF486302.1 GI:19744653
KEYWORDS
SOURCE       Human papillomavirus type 16
            Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
            Alphapapillomavirus.
            1 (bases 1 to 477)
AUTHORS      Chan, P.K.S., Lam, C.W., Cheung, T.H., Li, W.W.H., Lo, K.W.K.,
            Chan, M.Y.M., Cheung, J.L.K., Xu, L.Y. and Cheng, A.F.
TITLE        Human papillomavirus type 16 intraepithelial variant infection and risk
JOURNAL      Submitted (22-FEB-2002) Department of Microbiology, The Chinese
            University of Hong Kong, Prince of Wales Hospital, Shatin, N.T.,
            Hong Kong SAR, China

FEATURES     source
            1..477
            /organism="Human papillomavirus type 16"
            /mol_type="genomic DNA"
            /strain="As-P"
            /isolate="QEH-173"
            /db_xref="taxon:333760"
            /country="China"
            /note="Isolated from a patient with intraepithelial
            neoplasia grade I"

gene         1..477
            /gene="E6"
            /gene="E6"
            1..477
            /note="oncoprotein; transforming; trans-activating;
            p53-binding"
            /codon_start=1
            /product="E6 protein"

```

```

/strain="As-P"
/isolate="PWH-Q81"
/db_xref="taxon:333760"
/country="China"
/note="Isolated from a patient with invasive squamous cell
carcinoma of the cervix"

gene       1..477
            /gene="E6"
            /gene="E6"
            1..477
            /note="oncoprotein; transforming; trans-activating;
            p53-binding"
            /product="E6 protein"

ORIGIN
Query Match      100.0%; Score 28; DB 10; Length 477;
Best Local Similarity 100.0%; Pred. No. 0.031;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 GACATTATTGTTATAGTTGTATGGAAC 28
        |||||||
        251 GACATTATTGTTATAGTTGTATGGAAC 278

RESULT 29
AF486303      477 bp      DNA      linear      VRL 12-AUG-2002
LOCUS        Human papillomavirus type 16 strain As-C131 isolate QEH-173 E6
DEFINITION   protein (E6) gene, complete cds.
ACCESSION    AF486303
VERSION      AF486303.1 GI:19744655
KEYWORDS
SOURCE       Human papillomavirus type 16
            Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
            Alphapapillomavirus.
            1 (bases 1 to 477)
AUTHORS      Chan, P.K.S., Lam, C.W., Cheung, T.H., Li, W.W.H., Lo, K.W.K.,
            Chan, M.Y.M., Cheung, J.L.K., Xu, L.Y. and Cheng, A.F.
TITLE        Human papillomavirus type 16 intraepithelial variant infection and risk
JOURNAL      Submitted (22-FEB-2002) Department of Microbiology, The Chinese
            University of Hong Kong, Prince of Wales Hospital, Shatin, N.T.,
            Hong Kong SAR, China

FEATURES     source
            1..477
            /organism="Human papillomavirus type 16"
            /mol_type="genomic DNA"
            /strain="As-C131"
            /isolate="QEH-173"
            /db_xref="taxon:333760"
            /country="China"
            /note="Isolated from a patient with intraepithelial
            neoplasia grade I"

gene         1..477
            /gene="E6"
            /gene="E6"
            1..477
            /note="oncoprotein; transforming; trans-activating;
            p53-binding"
            /codon_start=1
            /product="E6 protein"

```

/protein_id="AA196608.1"
/db_xref="GI:19744656"
/translation="MHQKRTAMFODPQRPRLPOLCTELQTHIIEILCEVCCKQOL
LRREYDPAFRDLCTIVRDGNPAVDCDLKFKYSIKISRYHCYSLGITTLBOQYNKP
LCDLLIRINCOKPLCPDEKORHDKORFHNIRGRWTCRMSCCRSRTRRQTL"

ORIGIN

Query Match 100.0%; Score 28; DB 10; Length 477;
Best Local Similarity 100.0%; Pred. No. 0.031;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GACATTATTGTTATGTTGATGGAAC 28
|||||
251 GACATTATTGTTATGTTGATGGAAC 278

Db

RESULT 30
AF486304 477 bp DNA linear VRL 12-AUG-2002
LOCUS Human papillomavirus type 16 strain As-C442 isolate PMH-Q73 E6
DEFINITION
ACCESSION AF486304
VERSION AF486304.1 GI:19744657
KEYWORDS
SOURCE Human papillomavirus type 16
ORGANISM Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
Alphapapillomavirus.
REFERENCE 1 (bases 1 to 477)
AUTHORS Chan, P.K.S., Lam, C.W., Cheung, T.H., Li, W.W.H., Lo, K.W.K.,
Chan, M.Y.M., Cheung, J.L.K., Xu, L.Y. and Cheng, A.F.
TITLE Human papillomavirus type 16 intratypic variant infection and risk
for cervical neoplasia in Southern China
J. Infect. Dis. 186 (5), 696-700 (2002)
2 (bases 1 to 477)
AUTHORS Chan, P.K.S., Lam, C.W., Cheung, T.H., Li, W.W.H., Lo, K.W.K.,
Chan, M.Y.M., Cheung, J.L.K., Xu, L.Y. and Cheng, A.F.
TITLE Direct Submission
JOURNAL Submitted (22-FEB-2002) Department of Microbiology, The Chinese
University of Hong Kong, Prince of Wales Hospital, Shatin, N.T.,
Hong Kong SAR, China
LOCATION/Qualifiers

FEATURES
source 1..477
/organism="Human papillomavirus type 16"
/mol_type="genomic DNA"
/strain="As-C442"
/isolate="PMH-Q73"
/db_xref="taxon:333760"
/country="China"
/note="Isolated from a patient with invasive squamous cell
carcinoma of the cervix"
1..477
/gene="E6"
1..477
/gene="E6"
/note="oncoprotein; transforming; trans-activating;
p53-binding"
/codon_start=1
/product="E6 protein"
/protein_id="AA196609.1"
/db_xref="GI:19744658"
/translation="MHQKRTAMFODPQRPRLPOLCTELQTHIIEILCEVCCKQOL
LRREYDPAFRDLCTIVRDGNPAVDCDLKFKYSIKISRYHCYSLGITTLBOQYNKP
LCDLLIRINCOKPLCPDEKORHDKORFHNIRGRWTCRMSCCRSRTRRQTL"

ORIGIN

Query Match 100.0%; Score 28; DB 10; Length 477;
Best Local Similarity 100.0%; Pred. No. 0.031;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GACATTATTGTTATGTTGATGGAAC 28
|||||
251 GACATTATTGTTATGTTGATGGAAC 278

Db

RESULT 31
AF486305 477 bp DNA linear VRL 12-AUG-2002
LOCUS Human papillomavirus type 16 strain As-C442 isolate PMH-Q49 E6
DEFINITION
ACCESSION AF486305
VERSION AF486305.1 GI:19744659
KEYWORDS
SOURCE Human papillomavirus type 16
ORGANISM Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
Alphapapillomavirus.
REFERENCE 1 (bases 1 to 477)
AUTHORS Chan, P.K.S., Lam, C.W., Cheung, T.H., Li, W.W.H., Lo, K.W.K.,
Chan, M.Y.M., Cheung, J.L.K., Xu, L.Y. and Cheng, A.F.
TITLE Human papillomavirus type 16 intratypic variant infection and risk
for cervical neoplasia in Southern China
J. Infect. Dis. 186 (5), 696-700 (2002)
2 (bases 1 to 477)
AUTHORS Chan, P.K.S., Lam, C.W., Cheung, T.H., Li, W.W.H., Lo, K.W.K.,
Chan, M.Y.M., Cheung, J.L.K., Xu, L.Y. and Cheng, A.F.
TITLE Direct Submission
JOURNAL Submitted (22-FEB-2002) Department of Microbiology, The Chinese
University of Hong Kong, Prince of Wales Hospital, Shatin, N.T.,
Hong Kong SAR, China
LOCATION/Qualifiers

FEATURES
source 1..477
/organism="Human papillomavirus type 16"
/mol_type="genomic DNA"
/strain="As-C442"
/isolate="PMH-Q49"
/db_xref="taxon:333760"
/country="China"
/note="Isolated from a patient with invasive squamous cell
carcinoma of the cervix"
1..477
/gene="E6"
1..477
/gene="E6"
/note="oncoprotein; transforming; trans-activating;
p53-binding"
/codon_start=1
/product="E6 protein"
/protein_id="AA196610.1"
/db_xref="GI:19744660"
/translation="MHQKRTAMFODPQRPRLPOLCTELQTHIIEILCEVCCKQOL
LRREYDPAFRDLCTIVRDGNPAVDCDLKFKYSIKISRYHCYSLGITTLBOQYNKP
LCDLLIRINCOKPLCPDEKORHDKORFHNIRGRWTCRMSCCRSRTRRQTL"

ORIGIN

Query Match 100.0%; Score 28; DB 10; Length 477;
Best Local Similarity 100.0%; Pred. No. 0.031;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GACATTATTGTTATGTTGATGGAAC 28
|||||
251 GACATTATTGTTATGTTGATGGAAC 278

Db

RESULT 32
AF486306 477 bp DNA linear VRL 12-AUG-2002
LOCUS Human papillomavirus type 16 strain As-C193 isolate QSH-94 E6
DEFINITION
ACCESSION AF486306
VERSION AF486306.1 GI:19744661
KEYWORDS
SOURCE Human papillomavirus type 16
ORGANISM Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
Alphapapillomavirus.

```

REFERENCE
AUTHORS      1 (bases 1 to 477)
              Chan, P.K.S., Lam, C.W., Cheung, T.H., Li, W.W.H., Lo, K.W.K.,
              Chan, M.Y.M., Cheung, J.L.K., Xu, L.Y. and Cheng, A.F.
TITLE        Human papillomavirus type 16 intratypic variant infection and risk
              for cervical neoplasia in Southern China
JOURNAL      J. Infect. Dis. 186 (5), 696-700 (2002)
REFERENCE    2 (bases 1 to 477)
AUTHORS      Chan, P.K.S., Lam, C.W., Cheung, T.H., Li, W.W.H., Lo, K.W.K.,
              Chan, M.Y.M., Cheung, J.L.K., Xu, L.Y. and Cheng, A.F.
TITLE        Direct Submission
JOURNAL      Submitted (22-FEB-2002) Department of Microbiology, The Chinese
              University of Hong Kong, Prince of Wales Hospital, Shatin, N.T.,
              Hong Kong SAR, China
FEATURES
source       1..477
              /organism="Human papillomavirus type 16"
              /mol_type="genomic DNA"
              /strain="As-C193"
              /isolate="QEH-94"
              /db_xref="taxon:333760"
              /country="China"
              /note="Isolated from a patient with cervical
              intraepithelial neoplasia grade II"
gene         1..477
              /gene="E6"
              /note="oncoprotein; transforming; trans-activating;
              p53-binding"
              /codon_start=1
              /product="E6 protein"
              /protein_id="AAL96611.1"
              /db_xref="GI:19744662"
              /translation="MHQKRTAMFQDPRPRKLPOLCTELQTHIIEICVCKQOL
              LRREYDPAFDLCTVARDGNPAVCDKLCIKYSISEPHYCSLYGTTLEQYNKP
              LCDLIRICNCKPLCPBEKORHLDKORFHNIRGRWGRCMSCCRSSTRRETOL"
ORIGIN
Query Match      100.0%; Score 28; DB 10; Length 477;
Best Local Similarity 100.0%; Pred. No. 0.031;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Q# 1 GACATTATGTTATAGTTGTATGGAAC 28
DB 251 GACATTATGTTATAGTTGTATGGAAC 278

RESULT 33
AF486307      477 bp DNA linear VRL 12-AUG-2002
LOCUS        Human papillomavirus type 16 strain As-G137 isolate QEH-683 B6
DEFINITION   protein (E6) gene, complete cds.
ACCESSION    AF486307
VERSION      AF486307.1 GI:19744663
KEYWORDS
SOURCE       Human papillomavirus type 16
              Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
              Alphapapillomavirus.
ORGANISM     1 (bases 1 to 477)
              Chan, P.K.S., Lam, C.W., Cheung, T.H., Li, W.W.H., Lo, K.W.K.,
              Chan, M.Y.M., Cheung, J.L.K., Xu, L.Y. and Cheng, A.F.
TITLE        Human papillomavirus type 16 intratypic variant infection and risk
              for cervical neoplasia in Southern China
JOURNAL      J. Infect. Dis. 186 (5), 696-700 (2002)
REFERENCE    2 (bases 1 to 477)
AUTHORS      Chan, P.K.S., Lam, C.W., Cheung, T.H., Li, W.W.H., Lo, K.W.K.,
              Chan, M.Y.M., Cheung, J.L.K., Xu, L.Y. and Cheng, A.F.
TITLE        Direct Submission
JOURNAL      Submitted (22-FEB-2002) Department of Microbiology, The Chinese
              University of Hong Kong, Prince of Wales Hospital, Shatin, N.T.,
              Hong Kong SAR, China
FEATURES
location/Qualifiers

```

```

source       1..477
              /organism="Human papillomavirus type 16"
              /mol_type="genomic DNA"
              /strain="As-G137"
              /isolate="QEH-683"
              /db_xref="taxon:333760"
              /country="China"
              /note="Isolated from a patient with normal cervical
              cytology"
gene         1..477
              /gene="E6"
              /note="oncoprotein; transforming; trans-activating;
              p53-binding"
              /codon_start=1
              /product="E6 protein"
              /protein_id="AAL96612.1"
              /db_xref="GI:19744664"
              /translation="MHQKRTAMFQDPRPRKLPOLCTELQTHIIEICVCKQOL
              LRREYDPAFDLCTVARDGNPAVCDKLCIKYSISEPHYCSLYGTTLEQYNKP
              LCDLIRICNCKPLCPBEKORHLDKORFHNIRGRWGRCMSCCRSSTRRETOL"
ORIGIN
Query Match      100.0%; Score 28; DB 10; Length 477;
Best Local Similarity 100.0%; Pred. No. 0.031;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Q# 1 GACATTATGTTATAGTTGTATGGAAC 28
DB 251 GACATTATGTTATAGTTGTATGGAAC 278

RESULT 34
AF486308      477 bp DNA linear VRL 12-AUG-2002
LOCUS        Human papillomavirus type 16 strain As-A178 isolate QEH-186 E6
DEFINITION   protein (E6) gene, complete cds.
ACCESSION    AF486308
VERSION      AF486308.1 GI:19744665
KEYWORDS
SOURCE       Human papillomavirus type 16
              Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
              Alphapapillomavirus.
ORGANISM     1 (bases 1 to 477)
              Chan, P.K.S., Lam, C.W., Cheung, T.H., Li, W.W.H., Lo, K.W.K.,
              Chan, M.Y.M., Cheung, J.L.K., Xu, L.Y. and Cheng, A.F.
TITLE        Human papillomavirus type 16 intratypic variant infection and risk
              for cervical neoplasia in Southern China
JOURNAL      J. Infect. Dis. 186 (5), 696-700 (2002)
REFERENCE    2 (bases 1 to 477)
AUTHORS      Chan, P.K.S., Lam, C.W., Cheung, T.H., Li, W.W.H., Lo, K.W.K.,
              Chan, M.Y.M., Cheung, J.L.K., Xu, L.Y. and Cheng, A.F.
TITLE        Direct Submission
JOURNAL      Submitted (22-FEB-2002) Department of Microbiology, The Chinese
              University of Hong Kong, Prince of Wales Hospital, Shatin, N.T.,
              Hong Kong SAR, China
FEATURES
location/Qualifiers

```

```

p53-binding"
/codon_start=1
/product="E6 protein"
/protein_id="AAL96613.1"
/db_xref="GI:19744666"
/translation="MHOKRTAMFODPOERPRKLPOLCTELQTTIHIIILBCVCKQOOL
LRREVDPAPFDLCIVRDGNPYAVCDCLKFKYSISYRHYCSLYGTTLEQYNNKP
LCDLIRINCQKPLCPBEKQRHLDKQRFNIRGRWGRSCCSSRSTRRETQL"

ORIGIN
Query Match      100.0%; Score 28; DB 10; Length 477;
Best Local Similarity 100.0%; Pred. No. 0.031;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 35
AF486309      477 bp   DNA      linear   VRL 12-AUG-2002
LOCUS         Human papillomavirus type 16 strain As-A267 isolate PWM-066 E6
DEFINITION    AF486309
ACCESSION     AF486309
VERSION       AF486309.1 GI:19744667
KEYWORDS
SOURCE        Human papillomavirus type 16
              Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
              Alphapapillomavirus.
              1 (bases 1 to 477)
              Chan, P.K.S., Lam, C.W., Cheung, T.H., Li, W.W.H., Lo, K.W.K.,
              Chan, M.Y.M., Cheung, J.L.K., Xu, L.Y. and Cheng, A.F.
              Human papillomavirus type 16 intra-typic infection and risk
              for cervical neoplasia in Southern China
              J. Infect. Dis. 186 (5), 696-700 (2002)
              2 (bases 1 to 477)
              Chan, P.K.S., Lam, C.W., Cheung, T.H., Li, W.W.H., Lo, K.W.K.,
              Chan, M.Y.M., Cheung, J.L.K., Xu, L.Y. and Cheng, A.F.
              Direct Submission
              Submitted (22-FEB-2002) Department of Microbiology, The Chinese
              University of Hong Kong, Prince of Wales Hospital, Shatin, N.T.,
              Hong Kong SAR, China
              Location/Qualifiers
              1..477
               /organism="Human papillomavirus type 16"
               /mol_type="genomic DNA"
               /strain="As-A267"
               /isolate="PWM-066"
               /db_xref="taxon:333760"
               /country="China"
               /note="Isolated from a patient with invasive squamous cell
               carcinoma of the cervix"
               1..477
               /gene="E6"
               1..477
               /gene="E6"
               /note="oncoprotein; transforming; trans-activating;
               p53-binding"
               /codon_start=1
               /product="E6 protein"
               /protein_id="AAL96614.1"
               /db_xref="GI:19744668"
               /translation="MHOKRTAMFODPOERPRKLPOLCTELQTTIHIIILBCVCKQOOL
               LRREVDPAPFDLCIVRDGNPYAVCDCLKFKYSISYRHYCSLYGTTLEQYNNKP
               LCDLIRINCQKPLCPBEKQRHLDKQRFNIRGRWGRSCCSSRSTRRETQL"

ORIGIN
Query Match      100.0%; Score 28; DB 10; Length 477;
Best Local Similarity 100.0%; Pred. No. 0.031;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

ORIGIN
Query Match      100.0%; Score 28; DB 10; Length 477;
Best Local Similarity 100.0%; Pred. No. 0.031;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 36
AF486310      477 bp   DNA      linear   VRL 12-AUG-2002
LOCUS         Human papillomavirus type 16 strain E-P isolate PWM-042 E6 protein
DEFINITION    AF486310
ACCESSION     AF486310
VERSION       AF486310.1 GI:19744669
KEYWORDS
SOURCE        Human papillomavirus type 16
              Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
              Alphapapillomavirus.
              1 (bases 1 to 477)
              Chan, P.K.S., Lam, C.W., Cheung, T.H., Li, W.W.H., Lo, K.W.K.,
              Chan, M.Y.M., Cheung, J.L.K., Xu, L.Y. and Cheng, A.F.
              Human papillomavirus type 16 intra-typic infection and risk
              for cervical neoplasia in Southern China
              J. Infect. Dis. 186 (5), 696-700 (2002)
              2 (bases 1 to 477)
              Chan, P.K.S., Lam, C.W., Cheung, T.H., Li, W.W.H., Lo, K.W.K.,
              Chan, M.Y.M., Cheung, J.L.K., Xu, L.Y. and Cheng, A.F.
              Direct Submission
              Submitted (22-FEB-2002) Department of Microbiology, The Chinese
              University of Hong Kong, Prince of Wales Hospital, Shatin, N.T.,
              Hong Kong SAR, China
              Location/Qualifiers
              1..477
               /organism="Human papillomavirus type 16"
               /mol_type="genomic DNA"
               /strain="E-P"
               /isolate="PWM-042"
               /db_xref="taxon:333760"
               /country="China"
               /note="Isolated from a patient with invasive squamous cell
               carcinoma of the cervix"
               1..477
               /gene="E6"
               1..477
               /gene="E6"
               /note="oncoprotein; transforming; trans-activating;
               p53-binding"
               /codon_start=1
               /product="E6 protein"
               /protein_id="AAL96615.1"
               /db_xref="GI:19744670"
               /translation="MHOKRTAMFODPOERPRKLPOLCTELQTTIHIIILBCVCKQOOL
               LRREVDPAPFDLCIVRDGNPYAVCDCLKFKYSISYRHYCSLYGTTLEQYNNKP
               LCDLIRINCQKPLCPBEKQRHLDKQRFNIRGRWGRSCCSSRSTRRETQL"

ORIGIN
Query Match      100.0%; Score 28; DB 10; Length 477;
Best Local Similarity 100.0%; Pred. No. 0.031;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 37
AF486311      477 bp   DNA      linear   VRL 12-AUG-2002
LOCUS         Human papillomavirus type 16 strain B-P isolate PWM-Q15 E6 protein
DEFINITION    AF486311
ACCESSION     AF486311
VERSION       AF486311.1 GI:19744671
KEYWORDS
SOURCE        Human papillomavirus type 16
```


ORGANISM Human papillomavirus type 16
Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
Alphapapillomavirus.
REFERENCE 1 (bases 1 to 477)
AUTHORS Chan, P.K.S., Lam, C.W., Cheung, T.H., Li, W.W.H., Lo, K.W.K.,
Chan, M.Y.M., Cheung, J.L.K., Xu, L.Y., and Cheng, A.F.
TITLE Human papillomavirus type 16 intraepithelial neoplasia and risk
for cervical neoplasia in Southern China
JOURNAL J. Infect. Dis. 186 (5), 696-700 (2002)
REFERENCE 2 (bases 1 to 477)
AUTHORS Chan, P.K.S., Lam, C.W., Cheung, T.H., Li, W.W.H., Lo, K.W.K.,
Chan, M.Y.M., Cheung, J.L.K., Xu, L.Y., and Cheng, A.F.
TITLE Direct Submission
JOURNAL Submitted (22-FEB-2002) Department of Microbiology, The Chinese
University of Hong Kong, Prince of Wales Hospital, Shatin, N.T.,
Hong Kong SAR, China
FEATURES
source 1..477
/organism="Human papillomavirus type 16"
/mol_type="genomic DNA"
/strain="E-P"
/isolate="PMH-Q15"
/db_xref="taxon:333760"
/country="China"
/note="Isolated from a patient with invasive squamous cell
carcinoma of the cervix"
1..477
/gene="E6"
1..477
/gene="E6"
/note="oncoprotein; transforming; trans-activating;
p53-binding"
/codon_start=1
/product="E6 protein"
/protein_id="AAL96616.1"
/db_xref="GI:19744672"
/translation="MHQRTAMFODPQERPRKLPCLTETLTHDIILCEVCCKQOL
LRREYDPAFDLCIVRDGNPYAVCDKLTYSKISERHAYCSLYGTTLEQQYKPK
LCDLIRICNCKPFCPEBKQHLDKQRFNIRGRWTCRMSCCRSRTRETQL"
ORIGIN
Query Match 100.0%; Score 28; DB 10; Length 477;
Best Local Similarity 100.0%; Pred. No. 0.031;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GACATTATTTGTTATGTTGATGGAAC 28
Db 251 GACATTATTTGTTATGTTGATGGAAC 278
RESULT 38
AF486312 477 bp DNA linear VRL 12-AUG-2002
LOCUS Human papillomavirus type 16 strain E-P isolate QEH-525 E6 protein
DEFINITION (E6) gene, complete cds.
ACCESSION AF486312
VERSION AF486312.1 GI:19744673
KEYWORDS
SOURCE Human papillomavirus type 16
Organism: dsDNA viruses, no RNA stage; Papillomaviridae;
Alphapapillomavirus.
REFERENCE 1 (bases 1 to 477)
AUTHORS Chan, P.K.S., Lam, C.W., Cheung, T.H., Li, W.W.H., Lo, K.W.K.,
Chan, M.Y.M., Cheung, J.L.K., Xu, L.Y., and Cheng, A.F.
TITLE Human papillomavirus type 16 intraepithelial neoplasia and risk
for cervical neoplasia in Southern China
JOURNAL J. Infect. Dis. 186 (5), 696-700 (2002)
REFERENCE 2 (bases 1 to 477)
AUTHORS Chan, P.K.S., Lam, C.W., Cheung, T.H., Li, W.W.H., Lo, K.W.K.,
Chan, M.Y.M., Cheung, J.L.K., Xu, L.Y., and Cheng, A.F.
TITLE Direct Submission
JOURNAL Submitted (22-FEB-2002) Department of Microbiology, The Chinese

University of Hong Kong, Prince of Wales Hospital, Shatin, N.T.,
Hong Kong SAR, China
FEATURES
source 1..477
/organism="Human papillomavirus type 16"
/mol_type="genomic DNA"
/strain="E-P"
/isolate="QEH-525"
/db_xref="taxon:333760"
/country="China"
/note="Isolated from a patient with cervical
intraepithelial neoplasia grade II"
1..477
/gene="E6"
1..477
/gene="E6"
/note="oncoprotein; transforming; trans-activating;
p53-binding"
/codon_start=1
/product="E6 protein"
/protein_id="AAL96617.1"
/db_xref="GI:19744674"
/translation="MHQRTAMFODPQERPRKLPCLTETLTHDIILCEVCCKQOL
LRREYDPAFDLCIVRDGNPYAVCDKLTYSKISERHAYCSLYGTTLEQQYKPK
LCDLIRICNCKPFCPEBKQHLDKQRFNIRGRWTCRMSCCRSRTRETQL"
ORIGIN
Query Match 100.0%; Score 28; DB 10; Length 477;
Best Local Similarity 100.0%; Pred. No. 0.031;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GACATTATTTGTTATGTTGATGGAAC 28
Db 251 GACATTATTTGTTATGTTGATGGAAC 278
RESULT 39
AF486313 477 bp DNA linear VRL 12-AUG-2002
LOCUS Human papillomavirus type 16 strain E-P isolate QEH-453 E6 protein
DEFINITION (E6) gene, complete cds.
ACCESSION AF486313
VERSION AF486313.1 GI:19744675
KEYWORDS
SOURCE Human papillomavirus type 16
Organism: dsDNA viruses, no RNA stage; Papillomaviridae;
Alphapapillomavirus.
REFERENCE 1 (bases 1 to 477)
AUTHORS Chan, P.K.S., Lam, C.W., Cheung, T.H., Li, W.W.H., Lo, K.W.K.,
Chan, M.Y.M., Cheung, J.L.K., Xu, L.Y., and Cheng, A.F.
TITLE Human papillomavirus type 16 intraepithelial neoplasia and risk
for cervical neoplasia in Southern China
JOURNAL J. Infect. Dis. 186 (5), 696-700 (2002)
REFERENCE 2 (bases 1 to 477)
AUTHORS Chan, P.K.S., Lam, C.W., Cheung, T.H., Li, W.W.H., Lo, K.W.K.,
Chan, M.Y.M., Cheung, J.L.K., Xu, L.Y., and Cheng, A.F.
TITLE Direct Submission
JOURNAL Submitted (22-FEB-2002) Department of Microbiology, The Chinese
University of Hong Kong, Prince of Wales Hospital, Shatin, N.T.,
Hong Kong SAR, China
FEATURES
source 1..477
/organism="Human papillomavirus type 16"
/mol_type="genomic DNA"
/strain="E-P"
/isolate="QEH-453"
/db_xref="taxon:333760"
/country="China"
/note="Isolated from a patient with cervical
intraepithelial neoplasia grade III"
1..477
/gene="E6"

CDS

1..477
/gene="B6"
/note="Oncoprotein; transforming; trans-activating;
p53-binding"
/codon_start=1
/product="B6 protein"
/protein_id="AAL96618.1"
/db_xref="GI:19744678"
/translation="MHQKRTAMFQDPQRPRLPOLCTELQTHIDILIECVCKQOL
LRREYDPAFDLCTVYRDGNPAVCDKLCFKYSIRHYCYSLYGTTLQOQNK
LCDLILRCINQKPLCPBEKQRLDKQRFHNIIRGRTGRCMSCCRSSRTRETOL"

ORIGIN

Query Match 100.0%; Score 28; DB 10; Length 477;
Best Local Similarity 100.0%; Pred. No. 0.031;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 GACATTATGTTATAGTTGTATGGAAC 28
|||||
Db 251 GACATTATGTTATAGTTGTATGGAAC 278

RESULT 40
AF486315 477 bp DNA linear VRL 12-AUG-2002
LOCUS Human papillomavirus type 16 strain E-G276T isolate QEH-561 E6
DEFINITION Human papillomavirus type 16 strain E-G276T isolate QEH-561 E6
ACCESSION AF486315
VERSION AF486315.1 GI:19744679
KEYWORDS
SOURCE Human papillomavirus type 16
ORGANISM Human papillomavirus type 16
Virusess; deDNA virusess, no RNA stage; Papillomaviridae;
Alphapapillomavirus.
1 (bases 1 to 477)
REFERENCE Chan, P.K.S., Lam, C.W., Cheung, T.H., Li, W.W.H., Lo, K.W.K.,
AUTHORS Chan, M.Y.M., Cheung, J.L.K., Xu, L.Y. and Cheng, A.F.
TITLE Human papillomavirus type 16 intratypic variant infection and risk
JOURNAL for cervical neoplasia in Southern China
REFERENCES J. Infect. Dis. 186 (5), 696-700 (2002)
2 (bases 1 to 477)
AUTHORS Chan, P.K.S., Lam, C.W., Cheung, T.H., Li, W.W.H., Lo, K.W.K.,
Chan, M.Y.M., Cheung, J.L.K., Xu, L.Y. and Cheng, A.F.
TITLE Direct Submission
JOURNAL Submitted (22-FEB-2002) Department of Microbiology, The Chinese
AUTHORS University of Hong Kong, Prince of Wales Hospital, Shatin, N.T.,
Hong Kong SAR, China
LOCATION/Qualifiers
1..477
/organism="Human papillomavirus type 16"
/mol_type="genomic DNA"
/strain="E-G276T"
/isolate="QEH-561"
/db_xref="taxon:333760"
/country="China"
/note="Isolated from a patient with invasive squamous cell
carcinoma of the cervix"
1..477
/gene="B6"
1..477
/gene="B6"
/note="Oncoprotein; transforming; trans-activating;
p53-binding"
/codon_start=1
/product="B6 protein"
/protein_id="AAL96620.1"
/db_xref="GI:19744680"
/translation="MHQKRTAMFQDPQRPRLPOLCTELQTHIDILIECVCKQOL
LRREYDPAFDLCTVYRDGNPAVCDKLCFKYSIRHYCYSLYGTTLQOQNK
LCDLILRCINQKPLCPBEKQRLDKQRFHNIIRGRTGRCMSCCRSSRTRETOL"

ORIGIN

Query Match 100.0%; Score 28; DB 10; Length 477;

Best Local Similarity 100.0%; Pred. No. 0.031;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 GACATTATGTTATAGTTGTATGGAAC 28
|||||
Db 251 GACATTATGTTATAGTTGTATGGAAC 278

RESULT 41
AF486316 477 bp DNA linear VRL 12-AUG-2002
LOCUS Human papillomavirus type 16 strain E-G449T isolate QEH-1220 B6
DEFINITION Human papillomavirus type 16 strain E-G449T isolate QEH-1220 B6
ACCESSION AF486316
VERSION AF486316.1 GI:19744681
KEYWORDS
SOURCE Human papillomavirus type 16
ORGANISM Human papillomavirus type 16
Virusess; deDNA virusess, no RNA stage; Papillomaviridae;
Alphapapillomavirus.
1 (bases 1 to 477)
REFERENCE Chan, P.K.S., Lam, C.W., Cheung, T.H., Li, W.W.H., Lo, K.W.K.,
AUTHORS Chan, M.Y.M., Cheung, J.L.K., Xu, L.Y. and Cheng, A.F.
TITLE Human papillomavirus type 16 intratypic variant infection and risk
JOURNAL for cervical neoplasia in Southern China
REFERENCES J. Infect. Dis. 186 (5), 696-700 (2002)
2 (bases 1 to 477)
AUTHORS Chan, P.K.S., Lam, C.W., Cheung, T.H., Li, W.W.H., Lo, K.W.K.,
Chan, M.Y.M., Cheung, J.L.K., Xu, L.Y. and Cheng, A.F.
TITLE Direct Submission
JOURNAL Submitted (22-FEB-2002) Department of Microbiology, The Chinese
AUTHORS University of Hong Kong, Prince of Wales Hospital, Shatin, N.T.,
Hong Kong SAR, China
LOCATION/Qualifiers
1..477
/organism="Human papillomavirus type 16"
/mol_type="genomic DNA"
/strain="E-G449T"
/isolate="QEH-1220"
/db_xref="taxon:333760"
/country="China"
/note="Isolated from a patient with normal cervical
cytology"
1..477
/gene="B6"
1..477
/gene="B6"
/note="Oncoprotein; transforming; trans-activating;
p53-binding"
/codon_start=1
/product="B6 protein"
/protein_id="AAL96621.1"
/db_xref="GI:19744682"
/translation="MHQKRTAMFQDPQRPRLPOLCTELQTHIDILIECVCKQOL
LRREYDPAFDLCTVYRDGNPAVCDKLCFKYSIRHYCYSLYGTTLQOQNK
LCDLILRCINQKPLCPBEKQRLDKQRFHNIIRGRTGRCMSCCRSSRTRETOL"

ORIGIN

Query Match 100.0%; Score 28; DB 10; Length 477;
Best Local Similarity 100.0%; Pred. No. 0.031;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 GACATTATGTTATAGTTGTATGGAAC 28
|||||
Db 251 GACATTATGTTATAGTTGTATGGAAC 278

RESULT 42
AF486317 477 bp DNA linear VRL 12-AUG-2002
LOCUS Human papillomavirus type 16 strain E-G241T isolate QEH-1154 B6
DEFINITION Human papillomavirus type 16 strain E-G241T isolate QEH-1154 B6
ACCESSION AF486317

FEATURES	Chen, M.Y.M., Cheung, J.L.K., Xu, L.Y. and Cheng, A.F.
JOURNAL	Submitted (22-FEB-2002) Department of Microbiology, The Chinese University of Hong Kong, Prince of Wales Hospital, Shatin, N.T., Hong Kong SAR, China
SOURCE	Location/Qualifiers
	1..477
	/organism="Human papillomavirus type 16"
	/mol_type="genomic DNA"
	/strain="E-C442T"
	/isolate="QEH-196"
	/db_xref="taxon:333760"
	/country="China"
	/note="Isolated from a patient with normal cervical cytology"
gene	1..477
	/gene="E6"
	1..477
	/gene="E6"
CDS	/note="oncoprotein; transforming; trans-activating; p53-binding"
	/codon_start=1
	/product="E6 protein"
	/protein_id="AAL96623.1"
	/db_xref="GI:19744686"
	/translation="MHQKRTAMFQDPQERPRKLPLQCTELQTHIDILLCVCYCKQQLLRREVDPAFRDLCIVYRDNPYAVCDCKCPYKSKSEVHYCVSLYGLTLEQYNPKLCDLILRCINQCKPLCPDEKQRLHDKQRFHNIIRGWTGRCMSCRRFRRETQL"
ORIGIN	
Query Match	100.0%; Score 28; DB 10; Length 477;
Best Local Similarity	100.0%; Pred. No. 0.031;
Matches 28; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
Oy	1 GACATTATGTTAGTTGATGGAGC 28
Db	251 GACATTATGTTAGTTGATGGAGC 278
RESULT 44	
AF486319	477 bp DNA linear VRL 12-AUG-2002
LOCUS	Human papillomavirus type 16 strain E-7360T isolate QEH-352 B6
DEFINITION	Human papillomavirus type 16 strain E-7360T isolate QEH-352 B6
ACCESSION	AF486319
VERSION	AF486319
KEYWORDS	AF486319.1 GI:19744687
SOURCE	
ORGANISM	Human papillomavirus type 16
	Human papillomavirus type 16
	Viruses; dsDNA viruses, no RNA stage; Papillomaviridae; Alphapapillomavirus.
REFERENCE	1 (bases 1 to 477)
AUTHORS	Chan, P.K.S., Lam, C.W., Cheung, T.H., Li, W.W.H., Lo, K.W.K., Chan, M.Y.M., Cheung, J.L.K., Xu, L.Y. and Cheng, A.F.
TITLE	Human papillomavirus type 16 intracyclic variant infection and risk for cervical neoplasia in Southern China
JOURNAL	J. Infect. Dis. 186 (5), 696-700 (2002)
REFERENCE	2 (bases 1 to 477)
AUTHORS	Chan, P.K.S., Lam, C.W., Cheung, T.H., Li, W.W.H., Lo, K.W.K., Chan, M.Y.M., Cheung, J.L.K., Xu, L.Y. and Cheng, A.F.
TITLE	Direct Submmission
JOURNAL	Submitted (22-FEB-2002) Department of Microbiology, The Chinese University of Hong Kong, Prince of Wales Hospital, Shatin, N.T., Hong Kong SAR, China
FEATURES	Location/Qualifiers
SOURCE	1..477
	/organism="Human papillomavirus type 16"
	/mol_type="genomic DNA"
	/strain="E-7360T"
	/isolate="QEH-352"
	/db_xref="taxon:333760"
	/country="China"
	/note="Isolated from a patient with cervical

gene
1. .477
/gene="E6"
1. .477
/gene="E6"
/note="oncoprotein; transforming; trans-activating;
p53-binding"
/codon_start=1
/product="E6 protein"
/protein_id="AAL96624.1"
/db_xref="GI:19744688"
/translation="MHOKRTAMFODPOERPRKLPOLCTELQTTIHDIILCYCKQOL
LRBYVDPAFDLCIVRDGNPVACDCKLKYKISYRHYCYSLYGTLLDQYVNP
LCDLLIRICNCKPLCPBEKQRHLDKQRFNIRGRWGRCMSCCRSSRTRETLQ"

ORIGIN
Query Match 100.0%; Score 28; DB 10; Length 477;
Best Local Similarity 100.0%; Pred. No. 0.031;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACATTATTTGTTATGTTGATGGAAC 28
|||||
251 GACATTATTTGTTATGTTGATGGAAC 278
|||||

RESULT 45
AF486320 477 bp DNA linear VRL 12-AUG-2002
LOCUS Human papillomavirus type 16 strain E-G187T isolate QEH-400 E6
DEFINITION Human papillomavirus type 16 strain E-G187T isolate QEH-400 E6
ACCESSION AF486320
VERSION AF486320.1 GI:19744689
KEYWORDS
SOURCE Human papillomavirus type 16
ORGANISM Human papillomavirus type 16
Virusess; dsDNA viruses, no RNA stage; Papillomaviridae;
Alphapapillomavirus.
1 (bases 1 to 477)
Chan,P.K.S., Lam,C.W., Cheung,T.H., Li,M.W.H., Lo,K.W.K.,
Chan,M.Y.M., Cheung,J.L.K., Xu,L.Y. and Cheng,A.F.
Human papillomavirus type 16 intratypic variant infection and risk
for cervical neoplasia in Southern China
J. Infect. Dis. 186 (5), 696-700 (2002)
2 (bases 1 to 477)
Chan,P.K.S., Lam,C.W., Cheung,T.H., Li,M.W.H., Lo,K.W.K.,
Chan,M.Y.M., Cheung,J.L.K., Xu,L.Y. and Cheng,A.F.
Direct Submision
Submitted (22-FEB-2002) Department of Microbiology, The Chinese
University of Hong Kong, Prince of Wales Hospital, Shatin, N.T.,
Hong Kong SAR, China
Location/Qualifiers
1. .477
/organism="Human papillomavirus type 16"
/mol_type="genomic DNA"
/strain="E-G187T"
/isolate="QEH-400"
/db_xref="taxon:333760"
/country="China"
/note="Isolated from a patient with invasive squamous cell
carcinoma of the cervix"
1. .477
/gene="E6"
1. .477
/gene="E6"
/note="oncoprotein; transforming; trans-activating;
p53-binding"
/codon_start=1
/product="E6 protein"
/protein_id="AAL96625.1"
/db_xref="GI:19744690"
/translation="MHOKRTAMFODPOERPRKLPOLCTELQTTIHDIILCYCKQOL
LRBYVDPAFDLCIVRDGNPVACDCKLKYKISYRHYCYSLYGTLLDQYVNP
LCDLLIRICNCKPLCPBEKQRHLDKQRFNIRGRWGRCMSCCRSSRTRETLQ"

FEATURES
source
1. .477
/organism="Human papillomavirus type 16"
/mol_type="genomic DNA"
/strain="E-G187T"
/isolate="QEH-400"
/db_xref="taxon:333760"
/country="China"
/note="Isolated from a patient with invasive squamous cell
carcinoma of the cervix"
1. .477
/gene="E6"
1. .477
/gene="E6"
/note="oncoprotein; transforming; trans-activating;
p53-binding"
/codon_start=1
/product="E6 protein"
/protein_id="AAL96625.1"
/db_xref="GI:19744690"
/translation="MHOKRTAMFODPOERPRKLPOLCTELQTTIHDIILCYCKQOL
LRBYVDPAFDLCIVRDGNPVACDCKLKYKISYRHYCYSLYGTLLDQYVNP
LCDLLIRICNCKPLCPBEKQRHLDKQRFNIRGRWGRCMSCCRSSRTRETLQ"

ORIGIN
Query Match 100.0%; Score 28; DB 10; Length 477;
Best Local Similarity 100.0%; Pred. No. 0.031;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACATTATTTGTTATGTTGATGGAAC 28
|||||
251 GACATTATTTGTTATGTTGATGGAAC 278
|||||

RESULT 46
AF486321 477 bp DNA linear VRL 12-AUG-2002
LOCUS Human papillomavirus type 16 strain E-G538T isolate PMH-Q27 E6
DEFINITION Human papillomavirus type 16 strain E-G538T isolate PMH-Q27 E6
ACCESSION AF486321
VERSION AF486321.1 GI:19744691
KEYWORDS
SOURCE Human papillomavirus type 16
ORGANISM Human papillomavirus type 16
Virusess; dsDNA viruses, no RNA stage; Papillomaviridae;
Alphapapillomavirus.
1 (bases 1 to 477)
Chan,P.K.S., Lam,C.W., Cheung,T.H., Li,M.W.H., Lo,K.W.K.,
Chan,M.Y.M., Cheung,J.L.K., Xu,L.Y. and Cheng,A.F.
Human papillomavirus type 16 intratypic variant infection and risk
for cervical neoplasia in Southern China
J. Infect. Dis. 186 (5), 696-700 (2002)
2 (bases 1 to 477)
Chan,P.K.S., Lam,C.W., Cheung,T.H., Li,M.W.H., Lo,K.W.K.,
Chan,M.Y.M., Cheung,J.L.K., Xu,L.Y. and Cheng,A.F.
Direct Submision
Submitted (22-FEB-2002) Department of Microbiology, The Chinese
University of Hong Kong, Prince of Wales Hospital, Shatin, N.T.,
Hong Kong SAR, China
Location/Qualifiers
1. .477
/organism="Human papillomavirus type 16"
/mol_type="genomic DNA"
/strain="E-G538T"
/isolate="PMH-Q27"
/db_xref="taxon:333760"
/country="China"
/note="Isolated from a patient with invasive squamous cell
carcinoma of the cervix"
1. .477
/gene="E6"
1. .477
/gene="E6"
/note="oncoprotein; transforming; trans-activating;
p53-binding"
/codon_start=1
/product="E6 protein"
/protein_id="AAL96626.1"
/db_xref="GI:19744692"
/translation="MHOKRTAMFODPOERPRKLPOLCTELQTTIHDIILCYCKQOL
LRBYVDPAFDLCIVRDGNPVACDCKLKYKISYRHYCYSLYGTLLDQYVNP
LCDLLIRICNCKPLCPBEKQRHLDKQRFNIRGRWGRCMSCCRSSRTRETLQ"

ORIGIN
Query Match 100.0%; Score 28; DB 10; Length 477;
Best Local Similarity 100.0%; Pred. No. 0.031;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACATTATTTGTTATGTTGATGGAAC 28
|||||
251 GACATTATTTGTTATGTTGATGGAAC 278
|||||

RESULT 47
AF486323 477 bp DNA linear VRL 12-AUG-2002
LOCUS AF486323

DEFINITION Human papillomavirus type 16 strain E-G276G442T isolate QEH-2078 E6
ACCESSION AF486323
VERSION AF486323.1 GI:19744695
KEYWORDS
SOURCE Human papillomavirus type 16
ORGANISM Human papillomavirus type 16
REFERENCE 1 (bases 1 to 477)
AUTHORS Chan, P.K.S., Lam, C.W., Cheung, T.H., Li, W.W.H., Lo, K.W.K., Chan, M.Y.M., Cheung, J.L.K., Xu, L.Y., and Cheng, A.F.
TITLE Human papillomavirus type 16 intratypic variant infection and risk for cervical neoplasia in Southern China
JOURNAL J. Infect. Dis. 186 (5), 696-700 (2002)
REFERENCE 2 (bases 1 to 477)
AUTHORS Chan, P.K.S., Lam, C.W., Cheung, T.H., Li, W.W.H., Lo, K.W.K., Chan, M.Y.M., Cheung, J.L.K., Xu, L.Y., and Cheng, A.F.
TITLE Direct Submission
JOURNAL Submitted (22-FEB-2002) Department of Microbiology, The Chinese University of Hong Kong, Prince of Wales Hospital, Shatin, N.T., Hong Kong SAR, China
FEATURES
source location/Qualifiers
1..477
/organism="Human papillomavirus type 16"
/mol_type="genomic DNA"
/strain="E-G276G442T"
/isolate="QEH-2078"
/db_xref="taxon:333760"
/country="China"
/note="Isolated from a patient with cervical intraepithelial neoplasia grade II"
1..477
/gene="E6"
1..477
/gene="E6"
/note="oncoprotein; transforming; trans-activating; p53-binding"
/product="E6 protein"
/protein_id="AA196628.1"
/db_xref="GI:19744696"
/translation="MHQKATMPFODPOEPRRLPOLCTELQTHIDILIECVYCKQOLLRREYDPAFRLDCTIVRDGSPVACDCKIKYSIRYHRCYSLYGTLEQQNKPLCDLIRINCQKPLCPBEKQHLDKQRFHNIIRGRWCSCRSRRRTQL"
ORIGIN
Query Match 100.0%; Score 28; DB 10; Length 477;
Best Local Similarity 100.0%; Pred. No. 0.031;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GACATTATTGTTATAGTTGATGGAAC 28
Db 251 GACATTATTGTTATAGTTGATGGAAC 278
RESULT 48
AF404702 488 bp DNA linear VRL 09-OCT-2003
LOCUS Human papillomavirus type 16 isolate HPV16ECC11 E6 protein (E6)
DEFINITION Human papillomavirus type 16 isolate HPV16ECC11 E6 protein (E6)
ACCESSION AF404702
VERSION AF404702.1 GI:15529599
KEYWORDS
SOURCE Human papillomavirus type 16
ORGANISM Human papillomavirus type 16
REFERENCE 1 (bases 1 to 488)
AUTHORS Watts, K.J., Thompson, C.H., Cosgatch, Y.E. and Rose, B.R.
TITLE Sequence variation and physical state of human papillomavirus type 16 cervical cancer isolates from Australia and New Caledonia
JOURNAL Int. J. Cancer 97 (6), 868-874 (2002)

PUBMED 11857370
REFERENCE 2 (bases 1 to 488)
AUTHORS Watts, K.J., Thompson, C.H., Cosgatch, Y.E. and Rose, B.R.
TITLE Direct Submission
JOURNAL Submitted (31-JUL-2001) Department of Infectious Diseases, University of Sydney, Blackburn Building, D06, Off Western Avenue, Camperdown, Sydney, New South Wales 2006, Australia
FEATURES
source location/Qualifiers
1..488
/organism="Human papillomavirus type 16"
/mol_type="genomic DNA"
/isolate="HPV16ECC11"
/db_xref="taxon:333760"
1..432
/gene="E6"
1..432
/gene="E6"
/codon_start=1
/product="E6 protein"
/protein_id="AA101361.1"
/db_xref="GI:15529600"
/translation="PRKLPLQCTELQTHIDILIECVYCKQOLLRREYDPAFRLDCTIVRDGSPVACDCKIKYSIRYHRCYSLYGTLEQQNKPLCDLIRINCQKPLCPBEKQHLDKQRFHNIIRGRWCSCRSRRRTQL"
435..>488
/gene="E7"
435..>488
/gene="E7"
/codon_start=1
/product="E7 protein"
/protein_id="AA101362.1"
/db_xref="GI:15529601"
/translation="WHGDTPTLHEYMALDQPE"
ORIGIN
Query Match 100.0%; Score 28; DB 10; Length 488;
Best Local Similarity 100.0%; Pred. No. 0.031;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GACATTATTGTTATAGTTGATGGAAC 28
Db 206 GACATTATTGTTATAGTTGATGGAAC 233
RESULT 49
AR095319 519 bp DNA linear PAT 08-SEP-2000
LOCUS Human papillomavirus type 16 isolate HPV16ECC11 E6 protein (E6)
DEFINITION Human papillomavirus type 16 isolate HPV16ECC11 E6 protein (E6)
ACCESSION AR095319
VERSION AR095319.1 GI:10023090
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 519)
AUTHORS Edwards, S., John, C., Cox, J., Cooper, J., Webb, E., Ann, and Frazer, I.
TITLE Variants of human papillomavirus antigens
JOURNAL Patent: US 6004557-A 11 21-DEC-1999;
FEATURES
source location/Qualifiers
1..519
/organism="unknown"
/mol_type="unassigned DNA"
ORIGIN
Query Match 100.0%; Score 28; DB 2; Length 519;
Best Local Similarity 100.0%; Pred. No. 0.031;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GACATTATTGTTATAGTTGATGGAAC 28
Db 65 GACATTATTGTTATAGTTGATGGAAC 92

```

RESULT 50
AR095320      AR095320      519 bp      DNA      linear      PAT 08-SEP-2000
LOCUS
DEFINITION    Sequence 13 from patent US 6004557.
ACCESSION     AR095320
VERSION       AR095320.1  GI:10023092
KEYWORDS
SOURCE
ORGANISM      Unknown.
               Unclasseified.
REFERENCE     1 (bases 1 to 519)
AUTHORS      Edwards,S,John., Cox,J,Cooper., Webb,E,Ann. and Frazer,I.
TITLE        Variants of human papillomavirus antigens
JOURNAL      Patent: US 6004557-A 13 21 DEC-1999;
              Location/Qualifiers
FEATURES
             source
               /organism="unknown"
               /mol_type="unassigned DNA"

ORIGIN
Query Match          100.0%; Score 28; DB 2; Length 519;
Best Local Similarity 100.0%; Pred. No. 0.031;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  GACATTATTGTATAGTTGTATGGAAC 28
        |||
Db      458 GACATTATTGTATAGTTGTATGGAAC 485

```

Search completed: May 24, 2006, 06:52:46
 Job time : 926.046 secs

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 24, 2006, 05:46:37 ; Search time 851.046 Seconds
(without alignments)
2103.915 Million cell updates/sec

Title: US-10-601-913-5
Perfect score: 28
Sequence: 1 GAACAGCAATACAAACACCGTGTGTG 28

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 6366136 seqs, 31973710525 residues
Total number of hits satisfying chosen parameters: 12732272

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 1000 summaries

Database : GenEmbl : *
1: gb env : *
2: gb pat : *
3: gb db : *
4: gb pl : *
5: gb pr : *
6: gb ro : *
7: gb sts : *
8: gb ay : *
9: gb lun : *
10: gb vi : *
11: gb ov : *
12: gb htg : *
13: gb in : *
14: gb om : *
15: gb ba : *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	28	100.0	28	2 AR063359	AR063359 Sequence
2	28	100.0	28	2 AR349065	AR349065 Sequence
3	28	100.0	28	2 AR349066	AR349066 Sequence
4	28	100.0	28	2 AR349067	AR349067 Sequence
5	28	100.0	28	2 AR349068	AR349068 Sequence
6	28	100.0	87	2 AR577462	AR577462 Sequence
7	28	100.0	87	2 AX057044	AX057044 Sequence
8	28	100.0	245	10 HP014515	U14515 Human papil
9	28	100.0	253	10 HP014511	U14511 Human papil
10	28	100.0	271	10 HP014512	U14512 Human papil
11	28	100.0	272	10 HP014513	U14513 Human papil
12	28	100.0	273	10 HP014514	U14514 Human papil
13	28	100.0	274	10 HP014516	U14516 Human papil
14	28	100.0	421	10 S51110	S51110orf B6 [hum
15	28	100.0	451	10 AF404695	AF404695 Human pap
16	28	100.0	451	10 AF404696	AF404696 Human pap
17	28	100.0	451	10 AF404699	AF404699 Human pap
18	28	100.0	451	10 AF404700	AF404700 Human pap

19	28	100.0	451	10 AF404701	AF404701 Human pap
20	28	100.0	451	10 AF404703	AF404703 Human pap
21	28	100.0	456	2 AR167393	AR167393 Sequence
22	28	100.0	456	2 AR177943	AR177943 Sequence
23	28	100.0	456	10 AF327851	AF327851 Human pap
24	28	100.0	456	10 HP034107	HP034107 Human papil
25	28	100.0	456	10 HP034108	HP034108 Human papil
26	28	100.0	456	10 HP034109	HP034109 Human papil
27	28	100.0	456	10 HP034110	HP034110 Human papil
28	28	100.0	456	10 HP034111	HP034111 Human papil
29	28	100.0	456	10 HP034112	HP034112 Human papil
30	28	100.0	456	10 HP034113	HP034113 Human papil
31	28	100.0	456	10 HP034114	HP034114 Human papil
32	28	100.0	456	10 HP034115	HP034115 Human papil
33	28	100.0	456	10 HP034116	HP034116 Human papil
34	28	100.0	456	10 HP034117	HP034117 Human papil
35	28	100.0	456	10 HP034118	HP034118 Human papil
36	28	100.0	456	10 HP034119	HP034119 Human papil
37	28	100.0	456	10 HP034120	HP034120 Human papil
38	28	100.0	456	10 HP034121	HP034121 Human papil
39	28	100.0	456	10 HP034122	HP034122 Human papil
40	28	100.0	456	10 HP034123	HP034123 Human papil
41	28	100.0	456	10 HP034124	HP034124 Human papil
42	28	100.0	456	10 HP034125	HP034125 Human papil
43	28	100.0	456	10 HP034126	HP034126 Human papil
44	28	100.0	456	10 HP034127	HP034127 Human papil
45	28	100.0	456	10 HP034128	HP034128 Human papil
46	28	100.0	456	10 HP034129	HP034129 Human papil
47	28	100.0	456	10 HP034130	HP034130 Human papil
48	28	100.0	456	10 HP034131	HP034131 Human papil
49	28	100.0	456	10 HP034132	HP034132 Human papil
50	28	100.0	456	10 HP034133	HP034133 Human papil
51	28	100.0	456	10 HP034134	HP034134 Human papil
52	28	100.0	456	10 HP034135	HP034135 Human papil
53	28	100.0	458	2 AX658022	AX658022 Sequence
54	28	100.0	458	10 AF003013	AF003013 Human pap
55	28	100.0	458	10 AF003014	AF003014 Human pap
56	28	100.0	458	10 AF003015	AF003015 Human pap
57	28	100.0	458	10 AF003016	AF003016 Human pap
58	28	100.0	458	10 AF003017	AF003017 Human pap
59	28	100.0	458	10 AF003018	AF003018 Human pap
60	28	100.0	458	10 AF003019	AF003019 Human pap
61	28	100.0	473	10 AF404694	AF404694 Human pap
62	28	100.0	477	2 CS113240	CS113240 Sequence
63	28	100.0	477	2 AR577459	AR577459 Sequence
64	28	100.0	477	2 AX057038	AX057038 Sequence
65	28	100.0	477	2 AX766063	AX766063 Sequence
66	28	100.0	477	10 AF486299	AF486299 Human pap
67	28	100.0	477	10 AF486300	AF486300 Human pap
68	28	100.0	477	10 AF486301	AF486301 Human pap
69	28	100.0	477	10 AF486302	AF486302 Human pap
70	28	100.0	477	10 AF486303	AF486303 Human pap
71	28	100.0	477	10 AF486304	AF486304 Human pap
72	28	100.0	477	10 AF486305	AF486305 Human pap
73	28	100.0	477	10 AF486306	AF486306 Human pap
74	28	100.0	477	10 AF486307	AF486307 Human pap
75	28	100.0	477	10 AF486308	AF486308 Human pap
76	28	100.0	477	10 AF486309	AF486309 Human pap
77	28	100.0	477	10 AF486310	AF486310 Human pap
78	28	100.0	477	10 AF486311	AF486311 Human pap
79	28	100.0	477	10 AF486312	AF486312 Human pap
80	28	100.0	477	10 AF486313	AF486313 Human pap
81	28	100.0	477	10 AF486314	AF486314 Human pap
82	28	100.0	477	10 AF486315	AF486315 Human pap
83	28	100.0	477	10 AF486316	AF486316 Human pap
84	28	100.0	477	10 AF486317	AF486317 Human pap
85	28	100.0	477	10 AF486318	AF486318 Human pap
86	28	100.0	477	10 AF486319	AF486319 Human pap
87	28	100.0	477	10 AF486320	AF486320 Human pap
88	28	100.0	477	10 AF486321	AF486321 Human pap
89	28	100.0	477	10 AF486322	AF486322 Human pap
90	28	100.0	477	10 AF486323	AF486323 Human pap
91	28	100.0	477	10 AF486324	AF486324 Human pap

92	28	100.0	477	10	AF486325	AF486325 Human pap	165	28	100.0	7902	2	AX800450	AX800450 Sequence
93	28	100.0	477	10	DQ323982	DQ323982 Human pap	166	28	100.0	7904	2	BD070940	BD070940 Tissue sp
94	28	100.0	482	10	AF404698	AF404698 Human pap	167	28	100.0	7904	2	BD225188	BD225188 Human pap
95	28	100.0	488	10	AF404702	AF404702 Human pap	168	28	100.0	7904	2	CS073238	CS073238 Sequence
96	28	100.0	519	2	AR095319	AR095319 Sequence	169	28	100.0	7904	2	CS195904	CS195904 Sequence
97	28	100.0	519	2	AR173472	AR173472 Sequence	170	28	100.0	7904	2	I70142	I70142 Sequence
98	28	100.0	540	10	AF404705	AF404705 Human pap	171	28	100.0	7904	2	I70143	I70143 Sequence
99	28	100.0	543	10	AF404693	AF404693 Human pap	172	28	100.0	7904	2	I72345	I72345 Sequence
100	28	100.0	543	10	AF404697	AF404697 Human pap	173	28	100.0	7904	2	I72345	I72345 Sequence
101	28	100.0	543	10	AF404706	AF404706 Human pap	174	28	100.0	7904	2	AX763489	AX763489 Sequence
102	28	100.0	570	10	AY112663	AY112663 Human pap	175	28	100.0	7904	10	AX125673	AX125673 Sequence
103	28	100.0	592	10	AY098923	AY098923 Human pap	176	28	100.0	7904	10	AF472509	AF472509 Human pap
104	28	100.0	596	10	AY098919	AY098919 Human pap	177	28	100.0	7904	10	AF536179	AF536179 Human pap
105	28	100.0	600	10	AY098922	AY098922 Human pap	178	28	100.0	7905	10	PFH16	PFH16 Human pap
106	28	100.0	601	10	AY098918	AY098918 Human pap	179	28	100.0	7905	10	AF534061	AF534061 Human pap
107	28	100.0	602	10	AY112662	AY112662 Human pap	180	28	100.0	7906	10	HPH93348	HPH93348 Human pap
108	28	100.0	654	10	AF548023	AF548023 Human pap	181	28	100.0	7906	10	AR402678	AR402678 Human pap
109	28	100.0	693	10	AF469198	AF469198 Human pap	182	28	100.0	7906	10	AF536180	AF536180 Human pap
110	28	100.0	741	10	AF469197	AF469197 Human pap	183	28	100.0	7906	10	AY686579	AY686579 Human pap
111	28	100.0	755	10	AY089951	AY089951 Human pap	184	28	100.0	7906	10	AY686580	AY686580 Human pap
112	28	100.0	755	10	AY089952	AY089952 Human pap	185	28	100.0	7906	10	AY686581	AY686581 Human pap
113	28	100.0	755	10	AY089953	AY089953 Human pap	186	28	100.0	7907	10	AY686582	AY686582 Human pap
114	28	100.0	755	10	AY089954	AY089954 Human pap	187	28	100.0	7907	10	AY686583	AY686583 Human pap
115	28	100.0	755	10	AY089955	AY089955 Human pap	188	28	100.0	7908	10	AF472508	AF472508 Human pap
116	28	100.0	776	2	I19491	I19491 Sequence	189	28	100.0	7908	10	AF472508	AF472508 Human pap
117	28	100.0	779	2	AR199235	AR199235 Sequence	190	26.4	94.3	89	2	DD187313	DD187313 SPLICROSO
118	28	100.0	780	10	AF187866	AF187866 Human pap	191	26.4	94.3	99	2	DD187320	DD187320 SPLICROSO
119	28	100.0	780	10	AF187867	AF187867 Human pap	192	26.4	94.3	104	2	DD187318	DD187318 SPLICROSO
120	28	100.0	780	10	AF187868	AF187868 Human pap	193	26.4	94.3	104	2	DD187316	DD187316 SPLICROSO
121	28	100.0	780	10	AF187869	AF187869 Human pap	194	26.4	94.3	138	2	DD187319	DD187319 SPLICROSO
122	28	100.0	790	2	I88813	I88813 Sequence	195	26.4	94.3	149	2	DD187314	DD187314 SPLICROSO
123	28	100.0	801	2	AR095318	AR095318 Sequence	196	26.4	94.3	298	10	AF404704	AF404704 Human pap
124	28	100.0	802	2	AR173471	AR173471 Sequence	197	26.4	94.3	370	10	AF404692	AF404692 Human pap
125	28	100.0	822	2	BD080312	BD080312 Sequence	198	26.4	94.3	370	10	DD187859	DD187859 METHOD FO
126	28	100.0	822	2	BD080312	BD080312 Sequence	199	26.4	94.3	370	10	DD187859	DD187859 METHOD FO
127	28	100.0	822	2	BD103161	BD103161 Sequence	200	26.4	94.3	370	10	DD187859	DD187859 METHOD FO
128	28	100.0	822	2	AR183573	AR183573 Sequence	201	26.4	94.3	370	10	DD187859	DD187859 METHOD FO
129	28	100.0	822	2	AX020926	AX020926 Sequence	202	26.4	94.3	370	10	DD187859	DD187859 METHOD FO
130	28	100.0	879	2	A98748	A98748 Sequence	203	26.4	94.3	370	10	DD187859	DD187859 METHOD FO
131	28	100.0	879	2	BD080315	BD080315 Sequence	204	26.4	94.3	370	10	DD187859	DD187859 METHOD FO
132	28	100.0	879	2	BD103164	BD103164 Sequence	205	26.4	94.3	370	10	DD187859	DD187859 METHOD FO
133	28	100.0	879	2	AR183576	AR183576 Sequence	206	26.4	94.3	370	10	DD187859	DD187859 METHOD FO
134	28	100.0	879	2	AX020932	AX020932 Sequence	207	26.4	94.3	370	10	DD187859	DD187859 METHOD FO
135	28	100.0	921	10	HPA388069	HPA388069 Human pap	208	26.4	94.3	370	10	DD187859	DD187859 METHOD FO
136	28	100.0	939	2	AX403959	AX403959 Sequence	209	26.4	94.3	370	10	DD187859	DD187859 METHOD FO
137	28	100.0	939	2	AX460903	AX460903 Sequence	210	26.4	94.3	370	10	DD187859	DD187859 METHOD FO
138	28	100.0	1000	2	A07622	A07622 Amplified	211	26.4	94.3	370	10	DD187859	DD187859 METHOD FO
139	28	100.0	1000	2	I18920	I18920 Sequence	212	26.4	94.3	370	10	DD187859	DD187859 METHOD FO
140	28	100.0	1116	2	A98744	A98744 Sequence	213	26.4	94.3	370	10	DD187859	DD187859 METHOD FO
141	28	100.0	1116	2	BD080313	BD080313 Sequence	214	26.4	94.3	370	10	DD187859	DD187859 METHOD FO
142	28	100.0	1116	2	BD103162	BD103162 Sequence	215	26.4	94.3	370	10	DD187859	DD187859 METHOD FO
143	28	100.0	1116	2	AR183574	AR183574 Sequence	216	26.4	94.3	370	10	DD187859	DD187859 METHOD FO
144	28	100.0	1116	2	AX020928	AX020928 Sequence	217	26.4	94.3	370	10	DD187859	DD187859 METHOD FO
145	28	100.0	1173	2	A98752	A98752 Sequence	218	26.4	94.3	370	10	DD187859	DD187859 METHOD FO
146	28	100.0	1173	2	BD080317	BD080317 Sequence	219	26.4	94.3	370	10	DD187859	DD187859 METHOD FO
147	28	100.0	1173	2	BD103166	BD103166 Sequence	220	26.4	94.3	370	10	DD187859	DD187859 METHOD FO
148	28	100.0	1173	2	AR183578	AR183578 Sequence	221	26.4	94.3	370	10	DD187859	DD187859 METHOD FO
149	28	100.0	1173	2	AX020936	AX020936 Sequence	222	26.4	94.3	370	10	DD187859	DD187859 METHOD FO
150	28	100.0	1267	10	HPA242681	HPA242681 Human pap	223	26.4	94.3	370	10	DD187859	DD187859 METHOD FO
151	28	100.0	1267	10	HPA388056	HPA388056 Human pap	224	26.4	94.3	370	10	DD187859	DD187859 METHOD FO
152	28	100.0	1267	10	HPA388057	HPA388057 Human pap	225	26.4	94.3	370	10	DD187859	DD187859 METHOD FO
153	28	100.0	1267	10	HPA388058	HPA388058 Human pap	226	26.4	94.3	370	10	DD187859	DD187859 METHOD FO
154	28	100.0	1267	10	HPA388059	HPA388059 Human pap	227	26.4	94.3	370	10	DD187859	DD187859 METHOD FO
155	28	100.0	1267	10	HPA388060	HPA388060 Human pap	228	26.4	94.3	370	10	DD187859	DD187859 METHOD FO
156	28	100.0	1267	10	HPA388061	HPA388061 Human pap	229	26.4	94.3	370	10	DD187859	DD187859 METHOD FO
157	28	100.0	1267	10	HPA388062	HPA388062 Human pap	230	26.4	94.3	370	10	DD187859	DD187859 METHOD FO
158	28	100.0	1267	10	HPA388063	HPA388063 Human pap	231	26.4	94.3	370	10	DD187859	DD187859 METHOD FO
159	28	100.0	1267	10	HPA388064	HPA388064 Human pap	232	26.4	94.3	370	10	DD187859	DD187859 METHOD FO
160	28	100.0	1267	10	HPA388065	HPA388065 Human pap	233	26.4	94.3	370	10	DD187859	DD187859 METHOD FO
161	28	100.0	1267	10	HPA388066	HPA388066 Human pap	234	26.4	94.3	370	10	DD187859	DD187859 METHOD FO
162	28	100.0	1267	10	HPA388067	HPA388067 Human pap	235	26.4	94.3	370	10	DD187859	DD187859 METHOD FO
163	28	100.0	1267	10	HPA388068	HPA388068 Human pap	236	26.4	94.3	370	10	DD187859	DD187859 METHOD FO
164	28	100.0	3531	10	AF001599	AF001599 Human pap	237	26.4	94.3	370	10	DD187859	DD187859 METHOD FO

C 238	20	71.4	23	2	BD227561	BD227561 Assay usi	311	19	67.9	453	2	C0069257	C0069257 Sequence
C 239	20	71.4	23	2	AR476267	AR476267 Sequence	312	19	67.9	453	2	C0096287	C0096287 Sequence
C 240	20	71.4	24	2	BD017592	BD017592 Short-cha	313	19	67.9	453	2	C0020739	C0020739 Sequence
C 241	20	71.4	24	2	BD017593	BD017593 Short-cha	314	19	67.9	453	2	C0256922	C0256922 Sequence
C 242	20	71.4	24	2	117084	117084 Sequence 95	315	19	67.9	453	2	C0330879	C0330879 Sequence
C 243	20	71.4	24	2	117085	117085 Sequence 96	316	19	67.9	533	2	AX230117	AX230117 Sequence
C 244	20	71.4	51	2	DD178766	DD178766 METHOD FO	317	19	67.9	594	2	C0058669	C0058669 Sequence
C 245	20	71.4	51	2	AX802666	AX802666 Sequence	318	19	67.9	594	2	C0109032	C0109032 Sequence
C 246	20	71.4	51	2	AX802980	AX802980 Sequence	319	19	67.9	594	2	C0147663	C0147663 Sequence
C 247	20	71.4	73945	12	AC101549	AC101549 Mus muscu	320	19	67.9	594	2	C0207443	C0207443 Sequence
C 248	20	71.4	121416	12	CT573055	CT573055 Medicago	321	19	67.9	594	2	C0230911	C0230911 Sequence
C 249	20	71.4	124281	5	AL136101	AL136101 Human DNA	322	19	67.9	594	2	C0269059	C0269059 Sequence
C 250	20	71.4	126589	12	AC173645	AC173645 Strongylo	323	19	67.9	609	7	BV234777	BV234777 S234P6187
C 251	20	71.4	159035	5	AC092584	AC092584 Homo sapi	324	19	67.9	687	2	C0751610	C0751610 Sequence
C 252	20	71.4	163186	5	AC015994	AC015994 Homo sapi	325	19	67.9	803	10	PHR67	PHR67 Human papil
C 253	20	71.4	173452	5	AC003013	AC003013 Human PAC	326	19	67.9	924	2	C0078586	C0078586 Sequence
C 254	20	71.4	175482	6	AC132947	AC132947 Mus muscu	327	19	67.9	924	2	C0183596	C0183596 Sequence
C 255	20	71.4	176730	12	AC073994	AC073994 Homo sapi	328	19	67.9	924	2	C0207970	C0207970 Sequence
C 256	20	71.4	179884	6	AC121318	AC121318 Mus muscu	329	19	67.9	924	2	C0306749	C0306749 Sequence
C 257	20	71.4	243163	12	AC098002	AC098002 Rattus no	330	19	67.9	1089	2	C0821780	C0821780 Sequence
C 258	19.8	70.7	2671	11	BC072164	BC072164 Xenopus 1	331	19	67.9	1868	4	AP304366	AP304366 Rubus ida
C 259	19.8	70.7	185866	6	AC109510	AC109510 Mus muscu	332	19	67.9	1961	2	C0069411	C0069411 Sequence
C 260	19.8	70.7	213363	6	AC164399	AC164399 Mus muscu	333	19	67.9	1961	2	C0173728	C0173728 Sequence
C 261	19.8	70.7	225655	6	AC099690	AC099690 Mus muscu	334	19	67.9	1961	2	C0202870	C0202870 Sequence
C 262	19.6	70.0	551	7	BV462142	BV462142 r1c72h07.	335	19	67.9	1961	2	C0294250	C0294250 Sequence
C 263	19.6	70.0	641	7	BV602988	BV602988 S216P6054	336	19	67.9	2353	2	C0054455	C0054455 Sequence
C 264	19.6	70.0	831	7	BV494107	BV494107 S217P6475	337	19	67.9	2353	2	C0073696	C0073696 Sequence
C 265	19.6	70.0	1096	6	RATRBB1	L4158 Rattus norv	338	19	67.9	2353	2	C0104581	C0104581 Sequence
C 266	19.6	70.0	30000	2	DD164679	DD164679 NOVEL COM	339	19	67.9	2353	2	C0143291	C0143291 Sequence
C 267	19.6	70.0	30000	2	AF695899	AF695899 Sequence	340	19	67.9	2353	2	C0178780	C0178780 Sequence
C 268	19.6	70.0	68567	5	AL389916	AL389916 Noctoc sp	341	19	67.9	2353	2	C0203126	C0203126 Sequence
C 269	19.6	70.0	78229	5	AL139140	AL139140 Human DNA	342	19	67.9	2353	2	C0226485	C0226485 Sequence
C 270	19.6	70.0	108484	12	AC177279	AC177279 Strongylo	343	19	67.9	2353	2	C0264626	C0264626 Sequence
C 271	19.6	70.0	142650	5	AC112656	AC112656 Homo sapi	344	19	67.9	2353	2	C0301716	C0301716 Sequence
C 272	19.6	70.0	160431	12	AC118586	AC118586 Pan trogl	345	19	67.9	2353	2	C0338925	C0338925 Sequence
C 273	19.6	70.0	160431	12	AC024992	AC024992 Homo sapi	346	19	67.9	2720	5	HSB805169	HSB805169 Homo sapi
C 274	19.6	70.0	169336	12	AC073283	AC073283 Homo sapi	347	19	67.9	2811	5	AB220481	AB220481 Macaca fa
C 275	19.6	70.0	184220	12	AC173915	AC173915 Oryctolag	348	19	67.9	4432	5	AB039667	AB039667 Homo sapi
C 276	19.6	70.0	184220	12	AC019125	AC019125 Homo sapi	349	19	67.9	4432	5	BC112167	BC112167 Homo sapi
C 277	19.6	70.0	189203	12	AC153674	AC153674 Bos tauru	350	19	67.9	4888	2	C0720083	C0720083 Sequence
C 278	19.6	70.0	191983	5	AF274858	AF274858 Homo sapi	351	19	67.9	4895	2	CS056094	CS056094 Sequence
C 279	19.6	70.0	200059	12	AC155491	AC155491 Zea mays	352	19	67.9	4895	2	CS108411	CS108411 Sequence
C 280	19.6	70.0	211193	12	AC159760	AC159760 Bos tauru	353	19	67.9	4895	5	AF195657	AF195657 Homo sapi
C 281	19.6	70.0	230917	12	AC159760	AC159760 Bos tauru	354	19	67.9	7851	10	HPV35H	HPV35H Human papil
C 282	19.6	70.0	244251	12	AC125865	AC125865 Rattus no	355	19	67.9	8307	2	AX676351	AX676351 Sequence
C 283	19.6	70.0	247743	12	AC103091	AC103091 Rattus no	356	19	67.9	8307	15	BACWAPX	BACWAPX Bacillus su
C 284	19.6	70.0	290627	12	AC165991	AC165991 Bos tauru	357	19	67.9	8307	15	BACHYPP	BACHYPP Bacillus su
C 285	19.6	70.0	322475	13	U41276	U41276 Caenorhabd	358	19	67.9	17057	15	BACHUTWAPA	BACHUTWAPA
C 286	19.2	68.6	41039	13	AC160447	AC160447 Mus muscu	359	19	67.9	28954	12	AC101051	AC101051 Mus muscu
C 287	19.2	68.6	110000	15	CP000078_0	Continuation (30 o	360	19	67.9	35984	12	AC140910	AC140910
C 288	19.2	68.6	110000	15	CP000107_07	Continuation (8 o	361	19	67.9	35984	12	HSB6C3	HSB6C3
C 289	19.2	68.6	110000	15	AE016796_02	Continuation (3 o	362	19	67.9	36383	5	AF195658	AF195658
C 290	19.2	68.6	110000	15	BA000038_08	Continuation (9 o	363	19	67.9	36560	13	AF125956	AF125956
C 291	19.2	68.6	110000	15	AX950851_37	Continuation (38 o	364	19	67.9	38119	13	AY963685	AY963685
C 292	19.2	68.6	113845	6	AL935325	AL935325 Mouse DNA	365	19	67.9	38902	5	CBRG15A11	CBRG15A11
C 293	19.2	68.6	133845	6	AC138210	AC138210 Mus muscu	366	19	67.9	42688	13	AP001878	AP001878
C 294	19.2	68.6	206044	6	AC169668	AC169668 Mus muscu	367	19	67.9	46913	12	AP001878	AP001878
C 295	19.2	68.6	224712	12	AC105584	AC105584 Rattus no	368	19	67.9	67817	11	BX950227	BX950227 Zebrafish
C 296	19.2	68.6	240778	12	AC105584	AC105584 Rattus no	369	19	67.9	87054	11	BX511074	BX511074 Zebrafish
C 297	19	67.9	19	2	C0832282	C0832282 Sequence	370	19	67.9	94558	11	AX927353	AX927353 Zebrafish
C 298	19	67.9	19	2	C0880065	C0880065 Sequence	371	19	67.9	95528	5	AC096678	AC096678 Pan trogl
C 299	19	67.9	36	2	C0880065	C0880065 Sequence	372	19	67.9	103950	5	AC098123	AC098123
C 300	19	67.9	47	2	C0832283	C0832283 Sequence	373	19	67.9	108566	4	AC098123	AC098123
C 301	19	67.9	47	2	C0832283	C0832283 Sequence	374	19	67.9	110000	12	AP008212_214	AP008212_214
C 302	19	67.9	47	2	C0880066	C0880066 Sequence	375	19	67.9	110000	4	AE016959_214	AE016959_214
C 303	19	67.9	47	2	C0880066	C0880066 Sequence	376	19	67.9	110000	4	AE017349_00	AE017349_00
C 304	19	67.9	450	10	DD057309	DD057309 Human pap	377	19	67.9	110000	12	AC098192_2	AC098192_2
C 305	19	67.9	450	10	DD057310	DD057310 Human pap	378	19	67.9	110000	12	AC140833_3	AC140833_3
C 306	19	67.9	450	10	DD057311	DD057311 Human pap	379	19	67.9	110000	12	AP006488_2	AP006488_2
C 307	19	67.9	450	10	DD057312	DD057312 Human pap	380	19	67.9	110000	12	AC166314	AC166314
C 308	19	67.9	450	10	DD057313	DD057313 Human pap	381	19	67.9	115991	12	AY973543	AY973543
C 309	19	67.9	450	10	DD057314	DD057314 Human pap	382	19	67.9	120624	4	HSJ244P24	HSJ244P24
C 310	19	67.9	453	2	CQ054047	CQ054047 Sequence	383	19	67.9	125479	5		

384	19	67.9	129586	5	AC005086	AC005086 Homo sapi	457	19	67.9	207775	12	AC114183	AC114183 Rattus no
385	19	67.9	140228	11	BX248106	BX248106 Zebrafish	458	19	67.9	209164	11	CR589950	CR589950 Zebrafish
386	19	67.9	142525	12	AC146409	AC146409 Pan trogl	459	19	67.9	210210	12	AC173990	AC173990 Bos tauru
387	19	67.9	144139	11	BX321880	BX321880 Zebrafish	460	19	67.9	210437	12	CT571271	CT571271 Mus muscu
388	19	67.9	146006	5	AC005103	AC005103 Homo sapi	461	19	67.9	210856	12	AC106135	AC106135 Rattus no
389	19	67.9	146407	12	AC156689	AC156689 Bos tauru	462	19	67.9	211029	12	AC126840	AC126840 Rattus no
390	19	67.9	147117	4	AC069300	AC069300 Oryza sat	463	19	67.9	211209	12	CT485990	CT485990 Drosophill
391	19	67.9	148624	12	AC084815	AC084815 Homo sapi	464	19	67.9	211751	6	BX682545	BX682545 Mouse DNA
392	19	67.9	148758	12	AC176104	AC176104 Strongylo	465	19	67.9	217924	12	AC105529	AC105529 Rattus no
393	19	67.9	149387	6	AC173341	AC173341 Mus muscu	466	19	67.9	219879	12	AC097866	AC097866 Rattus no
394	19	67.9	149959	5	HSJ32324	HSJ32324 Human DNA	467	19	67.9	220808	12	AC018700	AC018700 Homo sapi
395	19	67.9	151964	5	AF047825	AF047825 Homo sapi	468	19	67.9	222383	12	AC178363	AC178363 Strongylo
396	19	67.9	152797	12	AC016856	AC016856 Homo sapi	469	19	67.9	222837	12	AC132028	AC132028 Rattus no
397	19	67.9	152842	12	AC067921	AC067921 Homo sapi	470	19	67.9	227437	12	AC119698	AC119698 Rattus no
398	19	67.9	153428	5	AC113387	AC113387 Homo sapi	471	19	67.9	228313	12	AC097557	AC097557 Rattus no
399	19	67.9	154730	13	AC008184	AC008184 Drosophill	472	19	67.9	231134	11	BX537109	BX537109 Zebrafish
400	19	67.9	156820	5	AC074323	AC074323 Homo sapi	473	19	67.9	232399	6	AC109258	AC109258 Mus muscu
401	19	67.9	158609	12	AC159121	AC159121 Bos tauru	474	19	67.9	232822	12	AC120204	AC120204 Pan trogl
402	19	67.9	158944	5	AC096746	AC096746 Homo sapi	475	19	67.9	232952	12	AC120963	AC120963 Rattus no
403	19	67.9	159431	6	AC128702	AC128702 Mus muscu	476	19	67.9	235286	5	AC136285	AC136285 Rattus no
404	19	67.9	160127	12	AC068977	AC068977 Homo sapi	477	19	67.9	235681	12	AC096454	AC096454 Rattus no
405	19	67.9	161002	12	CR759944	CR759944 Danio rer	478	19	67.9	238095	12	AC110424	AC110424 Rattus no
406	19	67.9	161523	5	AC012186	AC012186 Homo sapi	479	19	67.9	241394	12	AC131549	AC131549 Rattus no
407	19	67.9	163026	12	AC103883	AC103883 Homo sapi	480	19	67.9	243136	12	AC129786	AC129786 Rattus no
408	19	67.9	164440	6	AC157576	AC157576 Mus muscu	481	19	67.9	244254	12	AC168236	AC168236 Bos tauru
409	19	67.9	166108	12	AC114304	AC114304 Homo sapi	482	19	67.9	244759	12	AC098158	AC098158 Rattus no
410	19	67.9	166935	12	AC125957	AC125957 Rattus no	483	19	67.9	245977	12	AC127960	AC127960 Rattus no
411	19	67.9	166939	5	AL142643	AL142643 Homo sapi	484	19	67.9	247711	12	AC165494	AC165494 Bos tauru
412	19	67.9	167477	5	AC117402	AC117402 Homo sapi	485	19	67.9	252153	12	AC128765	AC128765 Rattus no
413	19	67.9	168015	11	BX005378	BX005378 Zebrafish	486	19	67.9	253645	12	AC137382	AC137382 Rattus no
414	19	67.9	168062	12	AC012857	AC012857 Drosophill	487	19	67.9	253895	12	AL772142	AL772142 Mus muscu
415	19	67.9	168623	6	AC159635	AC159635 Mus muscu	488	19	67.9	254463	12	AC117040	AC117040 Rattus no
416	19	67.9	169541	5	AL136178	AL136178 Human DNA	489	19	67.9	255501	6	AC105161	AC105161 Mus muscu
417	19	67.9	170709	12	AC021882	AC021882 Homo sapi	490	19	67.9	255943	12	AC109851	AC109851 Rattus no
418	19	67.9	171505	5	CNS01DIT5	AL137719 Human chr	491	19	67.9	258304	12	AC156060	AC156060 Bos tauru
419	19	67.9	172574	5	AC092667	AC092667 Homo sapi	492	19	67.9	258965	12	AC097783	AC097783 Rattus no
420	19	67.9	172942	6	AC016791	AC016791 Mus muscu	493	19	67.9	260027	13	AB003659	AB003659 Drosophill
421	19	67.9	173519	12	AC023800	AC023800 Homo sapi	494	19	67.9	261236	12	AC119381	AC119381 Rattus no
422	19	67.9	174217	12	AC021408	AC021408 Homo sapi	495	19	67.9	265674	12	AC087129	AC087129 Mus muscu
423	19	67.9	176012	12	AC011795	AC011795 Homo sapi	496	19	67.9	270667	12	AC129752	AC129752 Rattus no
424	19	67.9	176084	12	AC113354	AC113354 Homo sapi	497	19	67.9	299782	12	AC006844	AC006844 Caenorhab
425	19	67.9	176422	6	AC158170	AC158170 Mus muscu	498	19	67.9	310005	12	AC098428	AC098428 Rattus no
426	19	67.9	176578	12	AC176145	AC176145 Strongylo	499	19	67.9	334537	12	AC010187	AC010187 Homo sapi
427	19	67.9	176711	6	AC137969	AC137969 Mus muscu	500	19	67.9	337720	12	AC121050	AC121050 Rattus no
428	19	67.9	177324	11	BX937228	BX937228 Danio rer	501	18.8	67.1	70692	30	AR129428	AR129428 sequence
429	19	67.9	178231	12	CR293507	CR293507 Zebrafish	502	18.8	67.1	83000	5	AP007155_41	Continuation (42 o
430	19	67.9	178954	6	AC026963	AC026963 Homo sapi	503	18.8	67.1	110000	4	AB017343_11	Continuation (12 o
431	19	67.9	179875	12	BX004998	BX004998 Mouse DNA	504	18.8	67.1	110000	4	AP008209_219	Continuation (120
432	19	67.9	183591	12	AC135777	AC135777 Homo sapi	505	18.8	67.1	127210	12	AP006496_0	Continuation (220
433	19	67.9	184144	12	AC079734	AC079734 Mus muscu	506	18.8	67.1	127210	4	AC120537	AC120537 Oryza sat
434	19	67.9	189102	6	AC154646	AC154646 Mus muscu	507	18.8	67.1	153894	12	AC073233	AC073233 Homo sapi
435	19	67.9	189310	12	AC013419	AC013419 Homo sapi	508	18.8	67.1	166894	12	AC124794	AC124794 Homo sapi
436	19	67.9	191635	12	AC137760	AC137760 Homo sapi	509	18.8	67.1	166894	12	AC124794	AC124794 Homo sapi
437	19	67.9	192317	12	AC011986	AC011986 Homo sapi	510	18.8	67.1	166911	12	AC124795	AC124795 Homo sapi
438	19	67.9	192405	12	AC126160	AC126160 Rattus no	511	18.8	67.1	167357	4	AC109602	AC109602 Oryza sat
439	19	67.9	192653	6	AC134406	AC134406 Mus muscu	512	18.8	67.1	168252	5	AC079208	AC079208 Homo sapi
440	19	67.9	194020	6	AL611944	AL611944 Mouse DNA	513	18.8	67.1	175750	12	AC026297	AC026297 Homo sapi
441	19	67.9	194065	5	AC096679	AC096679 Homo sapi	514	18.8	67.1	205987	12	AC087488	AC087488 Homo sapi
442	19	67.9	194364	5	AC016820	AC016820 Homo sapi	515	18.8	67.1	208035	12	AC006520	AC006520 Mus muscu
443	19	67.9	194454	12	AC012357	AC012357 Homo sapi	516	18.8	67.1	210658	5	AC105204	AC105204 Homo sapi
444	19	67.9	195891	5	CNS00007	AL049780 Human chr	517	18.8	67.1	226607	12	CR847902	CR847902 Danio rer
445	19	67.9	197062	12	AC167365	AC167365 Bos tauru	518	18.6	66.4	128474	7	AC1018604	AC1018604 Homo sapi
446	19	67.9	199818	6	AL732563	AL732563 Mouse DNA	519	18.6	66.4	470	7	HS340107	HS340107 Homo sapi
447	19	67.9	199927	5	AC007055	AC007055 Homo sapi	520	18.6	66.4	577	7	BV348246	BV348246 S230P676R
448	19	67.9	200330	12	AC155097	AC155097 Bos tauru	521	18.6	66.4	652	7	BV575822	BV575822 GS91P6097
449	19	67.9	201486	12	AC091055	AC091055 Homo sapi	522	18.6	66.4	657	7	BV589453	BV589453 GS91P6265
450	19	67.9	203010	5	AC006530	AC006530 Homo sapi	523	18.6	66.4	724	7	BV589706	BV589706 GS91P6335
451	19	67.9	203084	5	AP004247	AP004247 Homo sapi	524	18.6	66.4	724	7	BV589706	BV589706 GS91P6405
452	19	67.9	203149	11	BX927330	BX927330 Zebrafish	525	18.6	66.4	962	7	BV587890	BV587890 GS91P6324
453	19	67.9	203901	15	BSUB0021	BSUB0021 Bacillus su	526	18.6	66.4	963	15	STPYKFORF	STPYKFORF X59945 S. cyphimuri
454	19	67.9	204439	12	AC129246	AC129246 Rattus no	527	18.6	66.4	8199	10	AY593803	AY593803 Foot-and-
455	19	67.9	204723	12	AC177927	AC177927 Zea mays	528	18.6	66.4	8217	10	AY593788	AY593788 Foot-and-
456	19	67.9	205790	12	AC128905	AC128905 Rattus no	529	18.6	66.4	21723	15	AB008760	AB008760 Salmonell

530	18.6	66.4	41039	4	T2D23	AC068143 Sequence	603	18.6	66.4	200882	12	AL591181
531	18.6	66.4	43707	5	AC112136	AC12136 Homo sapi	604	18.6	66.4	201088	5	AL589733
532	18.6	66.4	53610	5	AL589682	AL589682 Homo sapi	605	18.6	66.4	205457	12	AC159046
533	18.6	66.4	57904	12	AC091648	AC091648 Homo sapi	606	18.6	66.4	205087	12	AC179143
534	18.6	66.4	61169	5	AP001433	AP001433 Homo sapi	607	18.6	66.4	207021	6	AC101277
535	18.6	66.4	61812	5	AL591395	AL591395 Homo sapi	608	18.6	66.4	207359	12	AC174587
536	18.6	66.4	66311	5	AL591395	AL591395 Homo sapi	609	18.6	66.4	212060	12	AC148416
537	18.6	66.4	76795	12	AC023515	AC023515 Mouse DNA	610	18.6	66.4	213546	12	AC136602
538	18.6	66.4	78055	5	HS038361	HS038361 Homo sapi	611	18.6	66.4	214180	12	AL360082
539	18.6	66.4	80830	12	AC021244	AC021244 Homo sapi	612	18.6	66.4	215100	12	AC150972
540	18.6	66.4	82190	11	BSX08775	BSX08775 Zebrafish	613	18.6	66.4	216706	6	AL773534
541	18.6	66.4	87296	12	AC172793	AC172793 Bos taurus	614	18.6	66.4	220804	6	AC091283
542	18.6	66.4	96348	13	AC165428_7	AC165428 Homo sapi	615	18.6	66.4	222728	12	AC141289
543	18.6	66.4	100000	5	AP000158	AP000158 Homo sapi	616	18.6	66.4	227454	6	AC027700
544	18.6	66.4	110000	4	AP0007159_14	Continuation (15 o	617	18.6	66.4	227765	12	AC118099
545	18.6	66.4	110000	4	AP008207_353	Continuation (354	618	18.6	66.4	232165	12	AC098044
546	18.6	66.4	110000	12	HSY313F4_0	AL023808 Homo sapi	619	18.6	66.4	233050	15	AL627271
547	18.6	66.4	110000	12	HSY313F4_1	Continuation (2 of	620	18.6	66.4	234580	15	AL627271
548	18.6	66.4	110000	15	CP000230_24	Continuation (25 o	621	18.6	66.4	234949	12	AC168628
549	18.6	66.4	110000	15	CP000230_25	Continuation (26 o	622	18.6	66.4	241949	12	AC094309
550	18.6	66.4	110000	15	AB009951_08	Continuation (9 of	623	18.6	66.4	250515	12	AC167873
551	18.6	66.4	110000	15	AB009951_09	Continuation (10 o	624	18.6	66.4	280278	12	AL3390201
552	18.6	66.4	110000	15	AB014611_13	Continuation (14 o	625	18.6	66.4	281517	12	AC127144
553	18.6	66.4	110000	15	AB017220_14	Continuation (15 o	626	18.6	66.4	290973	12	AC093944
554	18.6	66.4	110000	15	CP000026_15	Continuation (16 o	627	18.6	66.4	300000	5	AP002530
555	18.6	66.4	110983	12	AC10549	AL079336 Homo sapi	628	18.6	66.4	302521	5	AB000014
556	18.6	66.4	117077	5	HS0662M14	CT573487 Dario rer	629	18.6	66.4	331211	5	AB016897
557	18.6	66.4	118325	12	CT573487	AL049760 Homo sapi	630	18.6	66.4	340000	5	AB001730
558	18.6	66.4	119118	5	HSR97C61	AL049760 Homo sapi	631	18.6	66.4	341469	12	AC111349
559	18.6	66.4	121647	4	AC135565	AC135565 Medicago	632	18.6	66.4	346203	12	BSX96407
560	18.6	66.4	122354	5	AL137251	AL137251 Homo sapi	633	18.4	65.7	320	7	HS3177C5
561	18.6	66.4	128241	5	AB041340	AB041340 Homo sapi	634	18.4	65.7	385	2	CO703959
562	18.6	66.4	139917	12	CT573112	CT573112 Dario rer	635	18.4	65.7	429	2	AX978295
563	18.6	66.4	141444	5	AP005015	AP005015 Homo sapi	636	18.4	65.7	429	2	AX978295
564	18.6	66.4	143621	5	AP001883	AP001883 Homo sapi	637	18.4	65.7	429	2	AX411601
565	18.6	66.4	146708	5	AC092834	AC092834 Homo sapi	638	18.4	65.7	477	8	AF373108
566	18.6	66.4	147340	12	CR854842	CR854842 Dario rer	639	18.4	65.7	542	2	AX544463
567	18.6	66.4	147392	11	CR753871	CR753871 Zebrafish	640	18.4	65.7	542	2	AX598636
568	18.6	66.4	149616	12	AC180034	AC180034 Strongylo	641	18.4	65.7	653	2	AX386377
569	18.6	66.4	151169	12	AC178287	AC178287 Strongylo	642	18.4	65.7	746	7	BV533228
570	18.6	66.4	154084	4	AP003734	AP003734 Oryza sat	643	18.4	65.7	785	7	BV511722
571	18.6	66.4	155629	5	AC125391	AC125391 Pan trogl	644	18.4	65.7	808	7	BV612979
572	18.6	66.4	156186	12	AC036150	AC036150 Homo sapi	645	18.4	65.7	2324	11	AB223056
573	18.6	66.4	159633	11	BX664714	BX664714 Zebrafish	646	18.4	65.7	2488	5	AK055765
574	18.6	66.4	162216	11	EX005461	EX005461 Zebrafish	647	18.4	65.7	2742	11	AY662635
575	18.6	66.4	163360	12	AC021598	AC021598 Homo sapi	648	18.4	65.7	2870	11	AY662637
576	18.6	66.4	163703	12	CT027785	CT027785 Dario rer	649	18.4	65.7	2879	11	AY662636
577	18.6	66.4	164766	12	AC018350	AC018350 Homo sapi	650	18.4	65.7	2920	11	AY662638
578	18.6	66.4	164834	6	AL929043	AL929043 Mouse DNA	651	18.4	65.7	3570	14	AF079767
579	18.6	66.4	166566	12	AC020746	AC020746 Homo sapi	652	18.4	65.7	5596	5	HSM805789
580	18.6	66.4	166598	12	AC170115	AC170115 Bos taurus	653	18.4	65.7	44059	5	AL668815
581	18.6	66.4	167127	5	AC097641	AC097641 Homo sapi	654	18.4	65.7	46864	12	AC100185
582	18.6	66.4	168768	5	AP006297	AP006297 Homo sapi	655	18.4	65.7	51924	5	AB009667
583	18.6	66.4	169806	12	AP007314	AP007314 Lotuss cor	656	18.4	65.7	52232	4	AB028617
584	18.6	66.4	170578	5	AP000789	AP000789 Homo sapi	657	18.4	65.7	57251	12	AC111150
585	18.6	66.4	172374	5	AB000924	AB000924 Homo sapi	658	18.4	65.7	64708	11	BX537144
586	18.6	66.4	173301	12	AC176630	AC176630 Strongylo	659	18.4	65.7	72363	12	AC101353
587	18.6	66.4	177962	6	AC147251	AC147251 Mus muscu	660	18.4	65.7	76385	5	AC007033
588	18.6	66.4	178619	6	AC152504	AC152504 Pan trogl	661	18.4	65.7	87161	5	AC109641
589	18.6	66.4	179306	11	BX072540	BX072540 Zebrafish	662	18.4	65.7	87616	12	AL133540
590	18.6	66.4	179306	6	AC154138	AC154138 Mus muscu	663	18.4	65.7	100204	5	HS076086
591	18.6	66.4	181302	5	AL355532	AL355532 Human DNA	664	18.4	65.7	102714	5	AP000692
592	18.6	66.4	182503	6	AL672195	AL672195 Mouse DNA	665	18.4	65.7	106148	5	HS73214
593	18.6	66.4	183285	6	AC128665	AC128665 Mus muscu	666	18.4	65.7	110000	4	AP007151_35
594	18.6	66.4	183488	6	AC142256	AC142256 Mus muscu	667	18.4	65.7	110000	4	AP008211_043
595	18.6	66.4	186557	12	CR936517	CR936517 Dario rer	668	18.4	65.7	110000	12	AC105546_0
596	18.6	66.4	189549	6	AC117190	AC117190 Mus muscu	669	18.4	65.7	110000	12	AC105546_0
597	18.6	66.4	190592	12	AC015806	AC015806 Homo sapi	670	18.4	65.7	110000	12	AC118422_0
598	18.6	66.4	193457	12	AC118103	AC118103 Rattus no	671	18.4	65.7	110000	15	CP000117_45
599	18.6	66.4	196520	6	AC111029	AC111029 Mus muscu	672	18.4	65.7	110000	15	CR954246_30
600	18.6	66.4	197027	11	AL929161	AL929161 Zebrafish	673	18.4	65.7	110000	15	AB016795_20
601	18.6	66.4	197864	5	HSY214H10	AL022344 Human DNA	674	18.4	65.7	110000	15	BA000019_28
602	18.6	66.4	199287	12	AC179373	AC179373 Strongylo	675	18.4	65.7	110000	15	BA000019_28

676	18.4	65.7	110000	15	BA000019_43	Continuation (44 o	749	18.4	65.7	199759	6	AP006505	AP006505 Mus muscu
677	18.4	65.7	110000	15	BA000019_44	Continuation (45 o	750	18.4	65.7	200000	5	AP000493	AP000493 Homo sapi
678	18.4	65.7	112361	5	AC023316	Homo sapi	751	18.4	65.7	202100	12	AC113218	AC113218 Rattus no
679	18.4	65.7	112670	5	AL589165	Homo sapi	752	18.4	65.7	203697	6	AC157906	AC157906 Mus muscu
680	18.4	65.7	113028	5	HS68D15	Homo sapi	753	18.4	65.7	204098	6	AC152176	AC152176 Mus muscu
681	18.4	65.7	113974	5	AC018763	Homo sapi	754	18.4	65.7	205222	6	AC123534	AC123534 Mus muscu
682	18.4	65.7	122839	5	AL161898	Human DNA	755	18.4	65.7	206252	12	AC141064	AC141064 Homo sapi
683	18.4	65.7	126753	6	AL1731655	Mouse DNA	756	18.4	65.7	206825	12	AC178933	AC178933 Zee maye
684	18.4	65.7	129098	5	HS46H23	Human DNA	757	18.4	65.7	207436	12	AC140823	AC140823 Homo sapi
685	18.4	65.7	132449	5	AL365272	Human DNA	758	18.4	65.7	207686	12	AC156812	AC156812 Bos tauru
686	18.4	65.7	133691	5	AC074347	Homo sapi	759	18.4	65.7	207757	12	AC166450	AC166450 Bos tauru
687	18.4	65.7	134245	5	HS1052M9	Human DNA	760	18.4	65.7	209271	6	AP006504	AP006504 Mus muscu
688	18.4	65.7	135299	5	AC091857	Homo sapi	761	18.4	65.7	209336	6	AC108840	AC108840 Mus muscu
689	18.4	65.7	137961	6	AL362990	Homo sapi	762	18.4	65.7	209999	6	AL731711	AL731711 Mouse DNA
690	18.4	65.7	139621	6	AL645745	Mouse DNA	763	18.4	65.7	210672	12	HS179115	HS179115 Homo sapi
691	18.4	65.7	138653	5	AL355381	Human DNA	764	18.4	65.7	211076	6	AC107611	AC107611 Rattus no
692	18.4	65.7	140549	5	AP006192	Homo sapi	765	18.4	65.7	211515	5	AC144479	AC144479 Pan trogl
693	18.4	65.7	141594	11	BX469928	Zebrafish	766	18.4	65.7	211662	6	AC161375	AC161375 Mus muscu
694	18.4	65.7	143167	12	AC182015	Atelerix	767	18.4	65.7	213050	15	AL646079	AL646079 Rattus no
695	18.4	65.7	143200	5	AC008413	Homo sapi	768	18.4	65.7	213950	12	AC135904	AC135904 Rattus no
696	18.4	65.7	144781	5	AC135426	Oryza sat	769	18.4	65.7	220212	12	AC141603	AC141603 Homo sapi
697	18.4	65.7	146687	5	AC110592	Homo sapi	770	18.4	65.7	220853	12	AC167604	AC167604 Bos tauru
698	18.4	65.7	147899	12	AC128194	Rattus no	771	18.4	65.7	222655	6	AC153541	AC153541 Mus muscu
699	18.4	65.7	152610	6	AC166051	Homo sapi	772	18.4	65.7	222880	12	AC172805	AC172805 Bos tauru
700	18.4	65.7	156630	12	AC177752	Strongylo	773	18.4	65.7	223110	6	AC123752	AC123752 Mus muscu
701	18.4	65.7	156967	6	AC145397	Rattus no	774	18.4	65.7	223536	12	CT027544	CT027544 Dantio xer
702	18.4	65.7	157034	12	AC125237	Homo sapi	775	18.4	65.7	224009	12	AC131058	AC131058 Mus muscu
703	18.4	65.7	158357	5	CNS07BPC	Human chr	776	18.4	65.7	224622	6	AC165953	AC165953 Mus muscu
704	18.4	65.7	159567	5	CNS07BPC	Human chr	777	18.4	65.7	228645	12	AC094273	AC094273 Rattus no
705	18.4	65.7	160177	11	BX363433	Zebrafish	778	18.4	65.7	229134	12	AC096974	AC096974 Rattus no
706	18.4	65.7	162058	11	BX530081	Zebrafish	779	18.4	65.7	229474	12	AC165510	AC165510 Bos tauru
707	18.4	65.7	162441	5	AL356321	Human DNA	780	18.4	65.7	230288	12	AC122977	AC122977 Rattus no
708	18.4	65.7	163283	5	AL355606	Human DNA	781	18.4	65.7	230786	12	AC114017	AC114017 Rattus no
709	18.4	65.7	164404	12	AC027149	Homo sapi	782	18.4	65.7	231378	11	AC146753	AC146753 Gallus ga
710	18.4	65.7	164605	12	AC113605	Homo sapi	783	18.4	65.7	232105	12	AC159779	AC159779 Bos tauru
711	18.4	65.7	164659	12	AC126577	Rattus no	784	18.4	65.7	232961	12	AC119317	AC119317 Rattus no
712	18.4	65.7	165235	6	AC145398	Rattus no	785	18.4	65.7	233038	12	AC118403	AC118403 Rattus no
713	18.4	65.7	166381	12	AC116943	Pan trogl	786	18.4	65.7	237729	12	AC098284	AC098284 Rattus no
714	18.4	65.7	166656	12	AC025331	Homo sapi	787	18.4	65.7	238988	12	AC096344	AC096344 Rattus no
715	18.4	65.7	169337	12	AC080110	Homo sapi	788	18.4	65.7	241544	12	AL713879	AL713879 Homo sapi
716	18.4	65.7	170398	6	AC154812	Mus muscu	789	18.4	65.7	244594	12	AC112470	AC112470 Rattus no
717	18.4	65.7	170398	12	AC013540	Homo sapi	790	18.4	65.7	245229	12	AC099093	AC099093 Rattus no
718	18.4	65.7	170425	11	BX465867	Zebrafish	791	18.4	65.7	245369	12	AC096044	AC096044 Rattus no
719	18.4	65.7	171255	5	AL691447	Human DNA	792	18.4	65.7	245617	12	AC095526	AC095526 Rattus no
720	18.4	65.7	172945	5	AC108099	Homo sapi	793	18.4	65.7	249765	12	AC129855	AC129855 Rattus no
721	18.4	65.7	173804	5	HS1030M6	Human DNA	794	18.4	65.7	251149	12	AC109774	AC109774 Rattus no
722	18.4	65.7	173945	11	BX322597	Zebrafish	795	18.4	65.7	252519	12	AC126215	AC126215 Rattus no
723	18.4	65.7	174085	6	AC166342	Mus muscu	796	18.4	65.7	253411	12	AC128606	AC128606 Rattus no
724	18.4	65.7	174600	6	AC110534	Mus muscu	797	18.4	65.7	253777	12	AC165904	AC165904 Bos tauru
725	18.4	65.7	174924	12	CT033833	Dantio xer	798	18.4	65.7	256974	12	AC103221	AC103221 Rattus no
726	18.4	65.7	176435	12	AC149874	Xenopus t	799	18.4	65.7	258862	6	CT009577	CT009577 Mouse DNA
727	18.4	65.7	176843	12	AC026655	Homo sapi	800	18.4	65.7	260224	12	AC130084	AC130084 Rattus no
728	18.4	65.7	177101	12	BX470137	Dantio xer	801	18.4	65.7	263706	12	AC097149	AC097149 Rattus no
729	18.4	65.7	178006	6	AC164068	Mus muscu	802	18.4	65.7	266718	12	AC098602	AC098602 Rattus no
730	18.4	65.7	178864	5	AC022535	Homo sapi	803	18.4	65.7	271368	12	AC170401	AC170401 Bos tauru
731	18.4	65.7	178956	6	AC122504	Mus muscu	804	18.4	65.7	278169	12	AC152273	AC152273 Bos tauru
732	18.4	65.7	180752	12	AC167845	Homo sapi	805	18.4	65.7	290052	12	AC160578	AC160578 Bos tauru
733	18.4	65.7	181501	12	AC016005	Homo sapi	806	18.4	65.7	301479	5	AC136590	AC136590 Homo sapi
734	18.4	65.7	183535	5	AC137551	Homo sapi	807	18.4	65.7	329752	12	AC105591	AC105591 Rattus no
735	18.4	65.7	183894	6	AC129935	Mus muscu	808	18.4	65.7	340000	5	AP001725	AP001725 Homo sapi
736	18.4	65.7	184510	6	AC101894	Homo sapi	809	18.4	65.0	131	2	CQ080978	CQ080978 Sequence
737	18.4	65.7	184876	5	AC084373	Homo sapi	810	18.2	65.0	131	2	CQ15461	CQ15461 Sequence
738	18.4	65.7	185121	6	AC102907	Mus muscu	811	18.2	65.0	131	2	CQ154273	CQ154273 Sequence
739	18.4	65.7	185920	5	AC009399	Homo sapi	812	18.2	65.0	131	2	CQ237518	CQ237518 Sequence
740	18.4	65.7	188920	12	AC140709	Homo sapi	813	18.2	65.0	131	2	CQ275142	CQ275142 Sequence
741	18.4	65.7	190613	12	AC011692	Homo sapi	814	18.2	65.0	131	2	CQ312116	CQ312116 Sequence
742	18.4	65.7	192707	12	AC116250	Rattus no	815	18.2	65.0	253	5	CQ349511	CQ349511 Sequence
743	18.4	65.7	195296	12	AC156652	Bos tauru	816	18.2	65.0	417	2	AX400859	AX400859 Sequence
744	18.4	65.7	195457	5	AC128064	Rattus no	817	18.2	65.0	514	2	CQ071773	CQ071773 Sequence
745	18.4	65.7	196470	12	BS000203	Pan trogl	818	18.2	65.0	514	2	CQ102392	CQ102392 Sequence
746	18.4	65.7	196715	12	AC141600	Homo sapi	819	18.2	65.0	514	2	CQ141323	CQ141323 Sequence
747	18.4	65.7	197153	6	AC141966	Rattus no	820	18.2	65.0	514	2	CQ224602	CQ224602 Sequence
748	18.4	65.7	199634	6	AL591209	Mouse DNA	821	18.2	65.0	514	2	CQ224602	CQ224602 Sequence

C 822	18.2	65.0	514	2	CQ262621	Sequence	895	18.2	65.0	178207	11	BX511263	BX511263 Zebrafish
C 823	18.2	65.0	514	2	CQ299678	Sequence	C 896	18.2	65.0	179207	12	AC009900	AC009900 Homo sapi
C 824	18.2	65.0	514	2	CQ336831	Sequence	897	18.2	65.0	187235	12	AP001785	AP001785 Homo sapi
C 825	18.2	65.0	522	8	AY890582	Synthetic	898	18.2	65.0	187891	12	AC109837	AC109837 Rattus no
C 826	18.2	65.0	522	8	AY893049	Synthetic	899	18.2	65.0	18875	12	AC171666	AC171666 Bos tauru
C 827	18.2	65.0	810	7	BV663693	S215P6580	C 900	18.2	65.0	200815	12	AC137450	AC137450 Rattus no
C 828	18.2	65.0	981	7	BV465234	G591P6126	901	18.2	65.0	202945	5	AC008038	AC008038 Homo sapi
C 829	18.2	65.0	1008	13	DQ062222	Lysiphleb	902	18.2	65.0	203078	12	AC083901	AC083901 Homo sapi
C 830	18.2	65.0	1084	13	AF052301	Drosophila	903	18.2	65.0	205372	12	AC117090	AC117090 Rattus no
C 831	18.2	65.0	1239	2	AR269521	Sequence	C 904	18.2	65.0	207891	6	AC153365	AC153365 Mus muscu
C 832	18.2	65.0	1275	2	BD003636	Disease a	905	18.2	65.0	214256	11	AC146680	AC146680 Gasterost
C 833	18.2	65.0	1550	4	BT000417	Arabidops	C 906	18.2	65.0	214606	11	AC161819	AC161819 Galus ga
C 834	18.2	65.0	1650	4	AY093049	Arabidops	C 907	18.2	65.0	217334	6	AC123695	AC123695 Mus muscu
C 835	18.2	65.0	1662	4	ATU04715	Arabidops	C 908	18.2	65.0	220482	12	AC112457	AC112457 Rattus no
C 836	18.2	65.0	1680	4	ATU06276	Arabidops	C 909	18.2	65.0	225868	6	AC098743	AC098743 Mus muscu
C 837	18.2	65.0	1695	4	AY086101	Arabidops	C 910	18.2	65.0	221985	12	AC156673	AC156673 Bos tauru
C 838	18.2	65.0	1706	4	ATM6731	Arabidops	911	18.2	65.0	234976	6	AC134623	AC134623 Mus muscu
C 839	18.2	65.0	1709	5	BC045571	X79210 A.thaliana	912	18.2	65.0	235453	12	AC131525	AC131525 Rattus no
C 840	18.2	65.0	1904	13	AK112351	Clona int	C 913	18.2	65.0	239924	12	AC112665	AC112665 Rattus no
C 841	18.2	65.0	2200	2	AX330718	Sequence	914	18.2	65.0	241340	12	AC126855	AC126855 Rattus no
C 842	18.2	65.0	2200	2	AX331139	Sequence	C 915	18.2	65.0	246199	12	AC119558	AC119558 Rattus no
C 843	18.2	65.0	2200	2	AX410722	Sequence	C 916	18.2	65.0	252527	12	AC114472	AC114472 Rattus no
C 844	18.2	65.0	2200	5	HS048296	Homo sapien	C 917	18.2	65.0	256294	2	CQ870382	CQ870382 Sequence
C 845	18.2	65.0	2208	2	CQ717029	Sequence	C 918	18.2	65.0	258545	12	AC114056	AC114056 Rattus no
C 846	18.2	65.0	2908	2	CQ717927	Sequence	C 919	18.2	65.0	262794	12	AC119702	AC119702 Rattus no
C 847	18.2	65.0	2916	2	AX925581	Sequence	C 920	18.2	65.0	278093	12	AC111376	AC111376 Rattus no
C 848	18.2	65.0	3387	5	BC023975	Homo sapi	C 921	18.2	65.0	308432	12	AC123212	AC123212 Rattus no
C 849	18.2	65.0	3918	4	ATU59737	Arabidops	C 922	18.2	65.0	313036	12	AC131879	AC131879 Rattus no
C 850	18.2	65.0	4342	2	BD224096	Sequence	923	18.2	64.3	313036	12	BD225198	BD225198 Human pap
C 851	18.2	65.0	4394	2	CQ827368	Sequence	924	18.2	64.3	313036	12	CS195914	CS195914 Sequence
C 852	18.2	65.0	5289	13	CQ414333	Sequence	925	18.2	64.3	313036	12	AX800464	AX800464 Sequence
C 853	18.2	65.0	6158	13	AF125957	Caenorhab	C 926	18.2	64.3	313036	12	AR146576	AR146576 Sequence
C 854	18.2	65.0	7933	11	QUNPMW	Japanese qu	C 927	18.2	64.3	313036	12	BD079825	BD079825 Sequence
C 855	18.2	65.0	11066	5	AF051160	Homo sapi	C 928	18.2	64.3	313036	12	AY487383	AY487383 Rattus no
C 856	18.2	65.0	48843	4	AP467900	Prunus pe	C 929	18.2	64.3	313036	12	BV394458	BV394458 Rattus no
C 857	18.2	65.0	64791	5	AL353649	Homo sapi	C 930	18.2	64.3	313036	12	BV222663	BV222663 Rattus no
C 858	18.2	65.0	73185	12	AC069580	Homo sapi	C 931	18.2	64.3	313036	12	BV192108	BV192108 Rattus no
C 859	18.2	65.0	91340	4	CT008898	Homo sapi	C 932	18.2	64.3	313036	12	BV023572	BV023572 Rattus no
C 860	18.2	65.0	107411	5	CT009490	M.truncat	933	18.2	64.3	313036	12	CR387333	CR387333 Gallus ga
C 861	18.2	65.0	110000	15	CR543861_29	Continuatio	934	18.2	64.3	313036	12	BV057860	BV057860 Rattus no
C 862	18.2	65.0	110000	15	CR543861_30	Continuatio	C 935	18.2	64.3	313036	12	BV624745	BV624745 Rattus no
C 863	18.2	65.0	110000	15	CR628336_04	Continuatio	C 936	18.2	64.3	313036	12	CR390519	CR390519 Gallus ga
C 864	18.2	65.0	110000	15	CR628337_04	Continuatio	937	18.2	64.3	313036	12	BX303662	BX303662 Gallus ga
C 865	18.2	65.0	110000	15	AB017354_04	Continuatio	938	18.2	64.3	313036	12	CR523066	CR523066 Gallus ga
C 866	18.2	65.0	110382	12	AC181107	Strongylo	939	18.2	64.3	313036	12	BX35940	BX35940 Gallus ga
C 867	18.2	65.0	117345	5	AP001775	Homo sapi	C 940	18.2	64.3	313036	12	CR524455	CR524455 Gallus ga
C 868	18.2	65.0	118439	5	AL135905	Homo sapi	C 941	18.2	64.3	313036	12	PU786462	PU786462 Polycha
C 869	18.2	65.0	119562	12	AC006779	Caenorhab	C 942	18.2	64.3	313036	12	AF251078	AF251078 Mus muscu
C 870	18.2	65.0	119942	4	AC007797	Arabidops	943	18.2	64.3	313036	12	AY206374	AY206374 Hepaticis
C 871	18.2	65.0	127562	5	AC087687	Homo sapi	C 944	18.2	64.3	313036	12	AY642231	AY642231 Plasmodiu
C 872	18.2	65.0	129049	12	AC069191	Homo sapi	C 945	18.2	64.3	313036	12	CO721578	CO721578 Sequence
C 873	18.2	65.0	132288	12	AC177394	Strongylo	C 946	18.2	64.3	313036	12	AF056618	AF056618 Homo sapi
C 874	18.2	65.0	138045	5	AC152477	Daelypus n	C 947	18.2	64.3	313036	12	AJ719672	AJ719672 Gallus ga
C 875	18.2	65.0	145035	12	CNS07ERY	Human chr	C 948	18.2	64.3	313036	12	AF216302	AF216302 Spodopter
C 876	18.2	65.0	146386	4	AC138130	Medicago	C 949	18.2	64.3	313036	12	MMU271753	MMU271753 Mus muscu
C 877	18.2	65.0	146736	5	AP001169	Homo sapi	950	18.2	64.3	313036	12	AC122112	AC122112 Bos tauru
C 878	18.2	65.0	151611	6	AC138596	Mus muscu	C 951	18.2	64.3	313036	12	CO577511	CO577511 Sequence
C 879	18.2	65.0	153278	12	AC040919	Homo sapi	C 952	18.2	64.3	313036	12	AL672202	AL672202 Mouse DNA
C 880	18.2	65.0	156166	12	AP001845	Homo sapi	C 953	18.2	64.3	313036	12	AB092808	AB092808 Locust cor
C 881	18.2	65.0	159279	12	AC174456	Strongylo	954	18.2	64.3	313036	12	AC133605	AC133605 Homo sapi
C 882	18.2	65.0	160096	12	AC022657	Homo sapi	955	18.2	64.3	313036	12	AC107646	AC107646 Mus muscu
C 883	18.2	65.0	160245	5	AC178215	Strongylo	C 956	18.2	64.3	313036	12	BX470143	BX470143 Zebrafish
C 884	18.2	65.0	161635	12	AC083873	Homo sapi	957	18.2	64.3	313036	12	AC101114	AC101114 Mus muscu
C 885	18.2	65.0	164671	12	AP002401	Homo sapi	C 958	18.2	64.3	313036	12	AC164712	AC164712 Bos tauru
C 886	18.2	65.0	167597	6	AC123551	Mus muscu	C 959	18.2	64.3	313036	12	AP005666	AP005666 Locust cor
C 887	18.2	65.0	167642	5	CNS07ERY	Human chr	C 960	18.2	64.3	313036	12	LC058083	LC058083 Locust cor
C 888	18.2	65.0	170348	12	AC021314	Homo sapi	961	18.2	64.3	313036	12	AC133605	AC133605 Homo sapi
C 889	18.2	65.0	171764	5	AL445225	Human DNA	962	18.2	64.3	313036	12	AC107646	AC107646 Mus muscu
C 890	18.2	65.0	174232	12	AC096251	Homo sapi	963	18.2	64.3	313036	12	BX470143	BX470143 Zebrafish
C 891	18.2	65.0	174770	12	AC177231	Strongylo	C 964	18.2	64.3	313036	12	AC154900	AC154900 Prunus pe
C 892	18.2	65.0	175037	12	AC180563	Strongylo	C 965	18.2	64.3	313036	12	AC007101	AC007101 Homo sapi
C 893	18.2	65.0	175643	12	AC018780	Homo sapi	966	18.2	64.3	313036	12	AC100537	AC100537 Mus muscu
C 894	18.2	65.0	177269	6	AC103632	Mus muscu	967	18.2	64.3	313036	12	AC096658	AC096658 Homo sapi

```

968 18 64.3 78428 12 AC174546 AC174546 Strongylo
c 969 18 64.3 89039 11 BX548062 BX548062 Zebrafish
970 18 64.3 100427 5 AC004048 AC004048 Homo sapi
c 971 18 64.3 108441 12 AC150445 AC150445 Medicago
c 972 18 64.3 109222 6 AC116371 AC116371 Mus muscu
c 973 18 64.3 109988 5 AC131151 AC131151 Homo sapi
c 974 18 64.3 110000 12 AC098940 AC098940 Rattus no
c 975 18 64.3 110000 12 CT005257 CT005257 no
976 18 64.3 110000 15 CP000091 CP000091 no
977 18 64.3 110000 15 CP000110 CP000110 no
978 18 64.3 110000 15 CP000117 CP000117 no
c 979 18 64.3 110000 15 CP000153 CP000153 no
c 980 18 64.3 114304 12 AC174308 AC174308 Medicago
c 981 18 64.3 118263 5 CENS01DRM CENS01DRM Homo chr
c 982 18 64.3 118684 5 AC025335 AC025335 Homo sapi
c 983 18 64.3 121017 5 AC087388 AC087388 Homo sapi
984 18 64.3 122057 12 AC177365 AC177365 Strongylo
c 985 18 64.3 127266 5 AC023672 AC023672 Homo sapi
c 986 18 64.3 128189 12 AC176261 AC176261 Strongylo
c 987 18 64.3 129477 11 BX927161 BX927161 Zebrafish
988 18 64.3 131365 4 AC140027 AC140027 Medicago
989 18 64.3 133485 6 AC118213 AC118213 Mus muscu
990 18 64.3 133819 12 AC178603 AC178603 Strongylo
c 991 18 64.3 134914 5 AC004542 AC004542 Homo sapi
c 992 18 64.3 135894 5 AC145425 AC145425 Homo sapi
c 993 18 64.3 136809 6 AL683892 AL683892 Mouse DNA
c 994 18 64.3 139077 11 AL935038 AL935038 Zebrafish
c 995 18 64.3 139342 10 AF325155 AF325155 Spodopter
c 996 18 64.3 139529 12 AC176009 AC176009 Strongylo
c 997 18 64.3 140222 11 CR385090 CR385090 Zebrafish
c 998 18 64.3 141146 12 AC179298 AC179298 Strongylo
c 999 18 64.3 142477 5 AC109781 AC109781 Homo sapi
1000 18

```

ALIGNMENTS

```

RESULT 1
AR063359 AR063359 28 bp DNA linear PAT 29-SEP-1999
DEFINITION Sequence 7 from patent US 5846701.
ACCESSION AR063359
VERSION AR063359.1 GI:5992667
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
1. .28
/mol_type="unknown"
/mol_type="unassigned DNA"
ORIGIN
Query Match 100.0%; Score 28; DB 2; Length 28;
Best Local Similarity 100.0%; Pred. No. 0.002;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 GAACAGCAATACAAACCGTTGTG 28
Db 1 GAACAGCAATACAAACCGTTGTG 28

```

```

RESULT 2
AR349065 AR349065 28 bp DNA linear PAT 17-AUG-2003
DEFINITION Sequence 5 from patent US 6583278.
ACCESSION AR349065

```

```

VERSION AR349065.1 GI:33749770
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
1. .28
/mol_type="unknown"
/mol_type="genomic DNA"
ORIGIN
Query Match 100.0%; Score 28; DB 2; Length 28;
Best Local Similarity 100.0%; Pred. No. 0.002;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 GAACAGCAATACAAACCGTTGTG 28
Db 1 GAACAGCAATACAAACCGTTGTG 28

```

```

RESULT 3
AR349066/c AR349066 28 bp DNA linear PAT 17-AUG-2003
DEFINITION Sequence 6 from patent US 6583278.
ACCESSION AR349066
VERSION AR349066.1 GI:33749771
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
1. .28
/mol_type="unknown"
/mol_type="genomic DNA"
ORIGIN
Query Match 100.0%; Score 28; DB 2; Length 28;
Best Local Similarity 100.0%; Pred. No. 0.002;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 GAACAGCAATACAAACCGTTGTG 28
Db 28 GAACAGCAATACAAACCGTTGTG 1

```

```

RESULT 4
AR349067 AR349067 28 bp DNA linear PAT 17-AUG-2003
DEFINITION Sequence 7 from patent US 6583278.
ACCESSION AR349067
VERSION AR349067.1 GI:33749772
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
1. .28
/mol_type="unknown"
/mol_type="genomic DNA"
ORIGIN
Query Match 100.0%; Score 28; DB 2; Length 28;
Best Local Similarity 100.0%; Pred. No. 0.002;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 GAACAGCAATACAAACCGTTGTG 28
Db 28 GAACAGCAATACAAACCGTTGTG 1

```

FEATURES Location/Qualifiers
source 1..28
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 100.0%; Score 28; DB 2; Length 28;
Best Local Similarity 100.0%; Pred. No. 0.002;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAACAGCAATACAAACCGTGTGTG 28
1 GAACAGCAATACAAACCGTGTGTG 28

Db 1 GAACAGCAATACAAACCGTGTGTG 28

RESULT 5
AR349068/c AR349068 28 bp DNA linear PAT 17-AUG-2003
LOCUS Sequence 8 from patent US 6583278.
DEFINITION AR349068
ACCESSION AR349068
VERSION AR349068.1 GI:33749773
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 28)
AUTHORS Carter,N.M.
TITLE Nucleic acid probes complementary to human papilloma virus nucleic acid
JOURNAL Patent: US 6583278-A 8 24-JUN-2003;
Gen-Probe Incorporated; San Diego, CA

FEATURES Location/Qualifiers
source 1..28
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 100.0%; Score 28; DB 2; Length 28;
Best Local Similarity 100.0%; Pred. No. 0.002;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAACAGCAATACAAACCGTGTGTG 28
28 GAACAGCAATACAAACCGTGTGTG 1

Db 28 GAACAGCAATACAAACCGTGTGTG 1

RESULT 6
AR577462 AR577462 87 bp DNA linear PAT 14-DEC-2004
LOCUS Sequence 7 from patent US 6783763.
DEFINITION AR577462
ACCESSION AR577462
VERSION AR577462.1 GI:56580067
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 87)
AUTHORS Choplin,J., Villada,I.B., Guillet,J.-G., Connan,F. and Ferries,E.
TITLE Polypeptidic proteolytic fragments of the E6 and E7 HPV proteins,
production and use thereof in vaccines
JOURNAL Patent: US 6783763-A 7 31-AUG-2004;
Peptide Immune Ligands and Institut National de la Sante et de la
Recherche Medicale (INSERM); Labège, Cedex.

FEATURES Location/Qualifiers
source 1..87
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 100.0%; Score 28; DB 2; Length 87;
Best Local Similarity 100.0%; Pred. No. 0.0022;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAACAGCAATACAAACCGTGTGTG 28
1 GAACAGCAATACAAACCGTGTGTG 28

Db 49 GAACAGCAATACAAACCGTGTGTG 76

RESULT 7
AX057044 AX057044 87 bp DNA linear PAT 17-JAN-2001
LOCUS Sequence 7 from Patent WO0075336.
DEFINITION AX057044
ACCESSION AX057044
VERSION AX057044.1 GI:12309885
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
REFERENCE 1
AUTHORS Choplin,J., Bourgault Villada,I., Guillet,J.G., Connan,F. and
Ferries,E.
TITLE Proteinfrafragmente die verschiedene epitope der hpv proteine e6 und
e7 umfassen, ihre herstellung und verwendungen insbesondere zur
impfung
JOURNAL Patent: WO 0075336-A 7 14-DEC-2000;
Biovector Therapeutics S.A. (FR) ; Institut National de la Sante et
de la Recherche Medicale (INSERM) (FR)

FEATURES Location/Qualifiers
source 1..87
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="fragment de la sequence codant pour E6 de HPV et
sequence peptidique correspondance"
1..>87
/note="unnamed protein product"
/codon_start=1
/transl_table=1
/protein_id="CAC22413.1"
/db_xref="GI:12309886"
/translation="MSERYRHYCVSLVGTITLQYKNRKLCDLLT"

ORIGIN
Query Match 100.0%; Score 28; DB 2; Length 87;
Best Local Similarity 100.0%; Pred. No. 0.0022;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAACAGCAATACAAACCGTGTGTG 28
1 GAACAGCAATACAAACCGTGTGTG 28

Db 49 GAACAGCAATACAAACCGTGTGTG 76

RESULT 8
HPV14515 HPV14515 245 bp DNA linear VRL 15-FEB-2001
LOCUS Human papillomavirus clone HPV1629 E6 gene, partial
DEFINITION HPV14515
ACCESSION HPV14515
VERSION HPV14515.1 GI:984951
KEYWORDS
SOURCE Human papillomavirus
ORGANISM Human papillomavirus
REFERENCE 1 (bases 1 to 245)
AUTHORS Haegert,D.G., Galutira,D.F. and Younghsband,B.H.
TITLE Sequence variation in the E6 gene of human papillomavirus type 16
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 245)
AUTHORS Galutira,D.F.
TITLE Direct Submission
JOURNAL Submitted (07-SEP-1994) Dante F. Galutira, Memorial University of
Newfoundland, Discipline of Pathology, Fac. Medicine, St. John's,
Newfoundland, A1B 3V6, Canada

FEATURES Location/Qualifiers

```

source
1. .245
/organism="Human papillomavirus"
/mol_type="genomic DNA"
/strain="HPV16"
/db_xref="caxon:10566"
/clone="HPV1629"
/tissue_type="cervical tissue from invasive carcinoma"

gene
1. .245
/gene="E6"
<1. .>245
/gene="E6"
/codon_start=2
/product="E6 protein"
/protein_id="AAB60569.2"
/db_xref="GI:1281650"
/translation="VDPAPRDLCIVRDGNPVAQCDKLFYSKISRYTCYVIG
TTTQQYNKPLCDLLIRCTNCPKPCPEEKORHDKK"

variation
61
/gene="E6"
/note="T to A transversion from PA16, GenBank Accession
Number K02718, position 286; NCBI gi: 333031"
/replace="t"

variation
64
/gene="E6"
/note="A to G transition from PA16, GenBank Accession
Number K02718, position 289; NCBI gi: 333031"
/replace="a"

variation
110
/gene="E6"
/note="C to T transition from PA16, GenBank Accession
Number K02718, position 335; NCBI gi: 333031"
/replace="c"

variation
125
/gene="E6"
/note="T to G transversion from PA16, GenBank Accession
Number K02718, position 350; NCBI gi: 333031"
/replace="t"

variation
139
/gene="E6"
/note="A to T transversion from PA16, GenBank Accession
Number K02718, position 364; NCBI gi: 333031"
/replace="a"

ORIGIN
Query Match 100.0%; Score 28; DB 10; Length 245;
Best Local Similarity 100.0%; Pred. No. 0.0024;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy
1 GAACGCAATACAAACAACCGTTGGTG 28
|||||
143 GAACGCAATACAAACAACCGTTGGTG 170

RESULT 9
HPV14511 253 bp DNA linear VRL 15-FEB-2001
LOCUS HPV14511
DEFINITION Human papillomavirus clone HPV1601 E6 protein (E6) gene, partial
cds.
ACCESSION U14511
VERSION U14511.1
KEYWORDS GI:984943
SOURCE Human papillomavirus
ORGANISM Human papillomavirus
Virusess; dsDNA viruses, no RNA stage; Papillomaviridae;
unclassified Papillomaviridae.
1 (bases 1 to 253)
Haegert,D.G., Galutira,D.F. and Younghuband,B.H.
Sequence variation in the E6 gene of human papillomavirus type 16
Unpublished
2 (bases 1 to 253)
Galutira,D.F.
Direct Submission
Submitted (07-SEP-1994) Dante F. Galutira, Memorial University of

```

```

FEATURES
    source
        Newfoundland, Discipline of Pathology, Fac. Medicine, St. John's,
        Newfoundland, A1B 3V6, Canada
        Location/Qualifiers
            1..253
                /organism="Human papillomavirus"
                /mol_type="genomic DNA"
                /strain="HPV16"
                /db_xref="taxon:10566"
                /clone="HPV1601"
                /cell_line="Caski and Stha"
                /tissue_type="cervical"
            1..253
                /gene="E6"
                /length=253
                /note="E6"
                /codon_start=1
                /product="E6 protein"
                /protein_id="AAB60565.2"
                /db_xref="GI:12831647"
                /translation="FAFRDLCLIVRDNPNPAVCDKCLKPYSKISBYRHYSVYGTTL
                EOOYNKPLCDLLIRTCINQKPLCPBEKQRLHDKQRPHNI"
            115
                /gene="E6"
                /note="T to G transversion from PA16, GenBank Accession
                Number K02718, position 350; NCBI gi: 333031"
                /replace="t"

    ORIGIN
        Query Match          100.0%; Score 28; DB 10; Length 253;
        Best Local Similarity 100.0%; Pred. No. 0.0024;
        Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

    Oy
        1 GAACGCAATACAAACAACCGTTGTGTG 28
          |||||
        133 GAACGCAATACAAACAACCGTTGTGTG 160

    Db
        133 GAACGCAATACAAACAACCGTTGTGTG 160

    RESULT 10
        HPV14512          271 bp DNA linear VRL 14-SEP-1995
        LOCUS
        DEFINITION
            Human papillomavirus clone HPV1603 E6 protein (E6) gene, partial
            cde.
        ACCESSION
            U14512
        VERSION
            U14512.1 GI:984945
        KEYWORDS
        SOURCE
            Human papillomavirus
            Human papillomavirus
            Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
            unclassified Papillomaviridae.
        ORGANISM
            1 (bases 1 to 271)
            Haegert,D.G., Galutira,D.F. and Youngsbend,B.H.
            Haegert,D.G., Galutira,D.F. and Youngsbend,B.H.
            Sequence variation in the E6 gene of human papillomavirus type 16
            2 (bases 1 to 271)
            Unpublished
            Galutira,D.F.
            Direct Submission
            Submitted (07-SEP-1994) Dante F. Galutira, Memorial University of
            Newfoundland, Discipline of Pathology, Fac. Medicine, St. John's,
            Newfoundland, A1B 3V6, Canada
            Location/Qualifiers
                1..271
                    /organism="Human papillomavirus"
                    /mol_type="genomic DNA"
                    /strain="HPV16"
                    /db_xref="taxon:10566"
                    /clone="HPV1603"
                    /tissue_type="cervical tissue from invasive carcinoma"
                1..271
                    /gene="E6"
                    /length=271
                    /note="E6"
                    /codon_start=2
                    /product="E6 protein"

```


variation
/protein_id="AAB60566.1"
/db_xref="GI:984946"
/translation="VYDPAFRLDCTIVYRDNPNVAVCDKCLFYKSISEYRHYCYSLYG
TTLEQQYNKPLCDLLIRNCINQKPLCEBEKQRLDKKQRFHNIRGR"
85
/gene="E6"
/note="T to G transversion from PA16, GenBank Accession
Number K02718, position 310; NCBI gi: 333031"
/replace="t"

ORIGIN
Query Match 100.0%; Score 28; DB 10; Length 271;
Best Local Similarity 100.0%; Pred. No. 0.0024;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAACAGCAATACAAACCGTGTGTG 28
143 GAACAGCAATACAAACCGTGTGTG 170

RESULT 11
HPV14513 272 bp DNA linear VRL 15-FEB-2001
LOCUS Human papillomavirus clone HPV1607 E6 protein (E6) gene, partial
DEFINITION cds.
ACCESSION U14513 GI:984947
VERSION U14513.1
KEYWORDS Human papillomavirus
SOURCE Human papillomavirus
ORGANISM Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
unclassified Papillomaviridae.
REFERENCE 1 (bases 1 to 272) Dante F. Galutira, D.F. and Younghusband, B.H.
Haegert, D.G., Galutira, D.F. and Younghusband, B.H.
TITLE Sequence variation in the E6 gene of human papillomavirus type 16
JOURNAL Unpublished
AUTHORS 2 (bases 1 to 272)
Galutira, D.F.
TITLE Direct Submission
AUTHORS Submitted (07-SEP-1994) Dante F. Galutira, Memorial University of
JOURNAL Newfoundland, Discipline of Pathology, Fac. Medicine, St. John's,
Newfoundland, A1B 3V6, Canada
LOCATION/Qualifiers
FEATURES
source
1..272
/organism="Human papillomavirus"
/mol_type="genomic DNA"
/strain="HPV16"
/db_xref="taxon:10566"
/clone="HPV1607"
/tissue_type="cervical tissue from invasive carcinoma"
1..272
/gene="E6"
/note="E6"
/length="272"
/gene="E6"
/product="E6 protein"
/codon_start=2
/protein_id="AAB60567.2"
/db_xref="GI:12831648"
/translation="VYDPAFRLDCTIVYRDNPNVAVCDKCLFYKSISEYRHYCYSLYG
TTLEQQYNKPLCDLLIRNCINQKPLCEBEKQRLDKKQRFHNIRGR"
31
/gene="E6"
/note="C to G transversion from PA16, GenBank Accession
Number K02718, position 256; NCBI gi: 333031"
/replace="c"

ORIGIN
Query Match 100.0%; Score 28; DB 10; Length 272;
Best Local Similarity 100.0%; Pred. No. 0.0024;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAACAGCAATACAAACCGTGTGTG 28
143 GAACAGCAATACAAACCGTGTGTG 28

Db 143 GAACAGCAATACAAACCGTGTGTG 170

RESULT 12
HPV14514 273 bp DNA linear VRL 15-FEB-2001
LOCUS Human papillomavirus clone HPV1627 E6 protein (E6) gene, partial
DEFINITION cds.
ACCESSION U14514 GI:984949
VERSION U14514.1
KEYWORDS Human papillomavirus
SOURCE Human papillomavirus
ORGANISM Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
unclassified Papillomaviridae.
REFERENCE 1 (bases 1 to 273)
Haegert, D.G., Galutira, D.F. and Younghusband, B.H.
TITLE Sequence variation in the E6 gene of human papillomavirus type 16
JOURNAL Unpublished
AUTHORS 2 (bases 1 to 273)
Galutira, D.F.
TITLE Direct Submission
AUTHORS Submitted (07-SEP-1994) Dante F. Galutira, Memorial University of
JOURNAL Newfoundland, Discipline of Pathology, Fac. Medicine, St. John's,
Newfoundland, A1B 3V6, Canada
LOCATION/Qualifiers
FEATURES
source
1..273
/organism="Human papillomavirus"
/mol_type="genomic DNA"
/strain="HPV16"
/db_xref="taxon:10566"
/clone="HPV1627"
/tissue_type="cervical tissue from invasive carcinoma"
1..273
/gene="E6"
/note="E6"
/length="273"
/gene="E6"
/product="E6 protein"
/codon_start=2
/protein_id="AAB60568.2"
/db_xref="GI:12831649"
/translation="VYDPAFRLDCTIVYRDNPNVAVCDKCLFYKSISEYRHYCYSLYG
TTLEQQYNKPLCDLLIRNCINQKPLCEBEKQRLDKKQRFHNIRGR"
61
/gene="E6"
/note="T to A transversion from PA16, GenBank Accession
Number K02718, position 286; NCBI gi: 333031"
/replace="t"
64
/gene="E6"
/note="A to G transition from PA16, GenBank Accession
Number K02718, position 289; NCBI gi: 333031"
/replace="a"
110
/gene="E6"
/note="C to T transition from PA16, GenBank Accession
Number K02718, position 335; NCBI gi: 333031"
/replace="c"

ORIGIN
Query Match 100.0%; Score 28; DB 10; Length 273;
Best Local Similarity 100.0%; Pred. No. 0.0024;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAACAGCAATACAAACCGTGTGTG 28
143 GAACAGCAATACAAACCGTGTGTG 170

Db 143 GAACAGCAATACAAACCGTGTGTG 170

RESULT 13
HPV14516 274 bp DNA linear VRL 14-SEP-1995
LOCUS Human papillomavirus clone HPV1649 E6 protein (E6) gene, partial
DEFINITION

ACCESSION U14516 GI:984953
VERSION U14516
KEYWORDS Human papillomavirus
SOURCE Human papillomavirus
ORGANISM Human papillomavirus
REFERENCE
AUTHORS unclassified Papillomaviridae.
TITLES 1 (bases 1 to 274)
JOURNAL Haeger, D.G., Galutira, D.F. and Youngnuband, B.H.
PUBMED Sequence variation in the E6 gene of human papillomavirus type 16
AUTHORS 2 (bases 1 to 274)
TITLES Galutira, D.F.
JOURNAL Direct Submission
REFERENCE Submitted (07-SEP-1994) Dante F. Galutira, Memorial University of
TITLES Newfoundland, Discipline of Pathology, Fac. Medicine, St. John's,
JOURNAL Newfoundland, A1B 3V6, Canada
FEATURES
source Location/Qualifiers
1..274
/organism="Human papillomavirus"
/mol_type="genomic DNA"
/strain="HPV16"
/db_xref="taxon:10566"
/clone="HPV1649"
/tissue_type="cervical tissue from invasive carcinoma"
1..274
/gene="E6"
/note="E6"
/codon_start=2
/product="E6 protein"
/protein_id="AAB60570.1"
/db_xref="GI:984954"
/translation="VYDPAFADLCIYRQGNPYAVCDKCKEYSKISEYRYCYSVYGT
TTTBQQYNKPLCDLIRINCQKPLCEBERKHLDDKKQPHNIRGM"
61
variation
/gene="E6"
/note="T to A transversion from PA16, GenBank Accession
Number K02718, position 286; NCBI gi: 333031"
/replace="c"
64
variation
/gene="E6"
/note="A to G transversion from PA16, GenBank Accession
Number K02718, position 289; NCBI gi: 333031"
/replace="a"
110
variation
/gene="E6"
/note="C to T transversion from PA16, GenBank Accession
Number K02718, position 335; NCBI gi: 333031"
/replace="c"
125
variation
/gene="E6"
/note="T to G transversion from PA16, GenBank Accession
Number K02718, position 350; NCBI gi: 333031"
/replace="c"
ORIGIN
Query Match 100.0%; Score 28; DB 10; Length 274;
Best Local Similarity 100.0%; Pred. No. 0.0024;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GAACAGCAATACAAACCGTTGTGTG 28
DB 143 GAACAGCAATACAAACCGTTGTGTG 170
RESULT 14
S51110 421 bp DNA linear VRL 08-MAY-1993
LOCUS orf E6 [human papillomavirus HPV, type 16, head and neck tumor,
DEFINITION Genomic, 421 nt].
ACCESSION S51110

VERSION S51110.1 GI:262061
KEYWORDS Human papillomavirus
SOURCE Human papillomavirus
ORGANISM Human papillomavirus
REFERENCE
AUTHORS unclassified Papillomaviridae.
TITLES 1 (bases 1 to 421)
JOURNAL Tyran, Y.S., Liu, S.T., Ong, W.R., Chen, M.L., Shu, C.H. and Chang, Y.S.
PUBMED Detection of Epstein-Barr virus and human papillomavirus in head
AUTHORS and neck tumors
JOURNAL J. Clin. Microbiol. 31 (1), 53-56 (1993)
PUBMED 8380183
REMARK Genbank staff at the National Library of Medicine created this
COMMENT entry [NCBI gisbgs 120826] from the original journal article.
FEATURES
source Location/Qualifiers
1..421
/organism="Human papillomavirus"
/mol_type="genomic DNA"
/db_xref="taxon:10566"
/note="type: 16"
1..421
/gene="orf E6"
ORIGIN
Query Match 100.0%; Score 28; DB 10; Length 421;
Best Local Similarity 100.0%; Pred. No. 0.0025;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GAACAGCAATACAAACCGTTGTGTG 28
DB 266 GAACAGCAATACAAACCGTTGTGTG 293
RESULT 15
AF404695 451 bp DNA linear VRL 09-OCT-2003
LOCUS Human papillomavirus type 16 isolate HPV16B6C4 E6 protein (E6) and
DEFINITION E7 protein (E7) genes, partial cds.
ACCESSION AF404695
VERSION AF404695.1 GI:15529578
KEYWORDS Human papillomavirus type 16
SOURCE Human papillomavirus type 16
ORGANISM Human papillomavirus
REFERENCE
AUTHORS unclassified Papillomaviridae.
TITLES 1 (bases 1 to 451)
JOURNAL Watts, K.J., Thompson, C.H., Coscort, Y.F. and Rose, B.R.
PUBMED Sequence variation and physical state of human papillomavirus type
AUTHORS 16 cervical cancer isolates from Australia and New Caledonia
JOURNAL Int. J. Cancer 97 (6), 868-874 (2002)
PUBMED 11857370
REFERENCE 2 (bases 1 to 451)
AUTHORS Watts, K.J., Thompson, C.H., Coscort, Y.F. and Rose, B.R.
TITLES Direct Submission
JOURNAL Submitted (31-JUL-2001) Department of Infectious Diseases,
University of Sydney, Blackburn Building, D06, Off Western Avenue,
Camperdown, Sydney, New South Wales 2006, Australia
FEATURES
source Location/Qualifiers
1..451
/organism="Human papillomavirus type 16"
/mol_type="genomic DNA"
/isolate="HPV16B6C4"
/db_xref="taxon:333760"
1..395
/gene="E6"
/note="E6"
1..395
/gene="E6"
/codon_start=3
/product="E6 protein"
/protein_id="AA01347.1"
/db_xref="GI:15529579"
/translation="TTHDITLTCVCKQQLLRREYVGFAPRDLCTYRQGNPYAVCDK

CLAFVSKISSEYRYCVSLVGTTLLEQYKNPCLDLLIRNCINQCKPLCPBEKQRHLDDKKO
RFINIRGRWTGRMCCSCRSSRTREITOL"
398..>451
/gene="E7"
/db_xref="GI:15529580"
/translation="MHGDTPTLHEMYMLDLPPE"

ORIGIN

Query Match 100.0%; Score 28; DB 10; Length 451;
Best Local Similarity 100.0%; Pred. No. 0.0025;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAACAGCAATACAAACCGTTGTGTG 28
|||||
204 GAACAGCAATACAAACCGTTGTGTG 231

Db

RESULT 16
AF404696 451 bp DNA linear VRL 09-OCT-2003
LOCUS Human papillomavirus type 16 isolate HVI6B6CC5 B6 protein (B6) and
DEFINITION E7 protein (E7) genes, partial cds.
ACCESSION AF404696
VERSION AF404696.1 GI:15529581
KEYWORDS Human papillomavirus type 16
SOURCE Human papillomavirus type 16
ORGANISM Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
Alphapapillomavirus.
REFERENCE 1 (bases 1 to 451)
AUTHORS Watts,K.J., Thompson,C.H., Cossart,Y.E. and Rose,B.R.
TITLE Sequence variation and physical state of human papillomavirus type
16 cervical cancer isolates from Australia and New Caledonia
JOURNAL Int. J. Cancer 97 (6), 868-874 (2002)
PUBMED 11857370
REFERENCE 2 (bases 1 to 451)
AUTHORS Watts,K.J., Thompson,C.H., Cossart,Y.E. and Rose,B.R.
TITLE Direct Submission
JOURNAL Submitted (31-JUL-2001) Department of Infectious Diseases,
University of Sydney, Blackburn Building, D06, Off Western Avenue,
Camperdown, Sydney, New South Wales 2006, Australia
Location/Qualifiers
1..451
/organism="Human papillomavirus type 16"
/mol_type="genomic DNA"
/isolate="HPV16B6CC5"
/db_xref="taxon:333760"
<1..395
/gene="E6"
<1..395
/gene="E6"
/codon_start=1
/product="E6 protein"
/protein_id="AAL01349.1"
/db_xref="GI:15529582"
/translation="TTHNIIIECVYCKQQLRREVDPAFRDLCTVYRDGNPYAVCDK
CLKFYSKISSEYRYCVSVGTTLLEQYKNPCLDLLIRNCINQCKPLCPBEKQRHLDDKKO
RFINIRGRWTGRMCCSCRSSRTREITOL"
398..>451
/gene="E7"
398..>451
/codon_start=1
/product="E7 protein"
/protein_id="AAL01350.1"
/db_xref="GI:15529583"
/translation="MHGDTPTLHEMYMLDLPPE"

ORIGIN

Query Match 100.0%; Score 28; DB 10; Length 451;
Best Local Similarity 100.0%; Pred. No. 0.0025;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAACAGCAATACAAACCGTTGTGTG 28
|||||
204 GAACAGCAATACAAACCGTTGTGTG 231

Db

RESULT 17
AF404699 451 bp DNA linear VRL 09-OCT-2003
LOCUS Human papillomavirus type 16 isolate HVI6B6CC8 B6 protein (B6) and
DEFINITION E7 protein (E7) genes, partial cds.
ACCESSION AF404699
VERSION AF404699.1 GI:15529590
KEYWORDS Human papillomavirus type 16
SOURCE Human papillomavirus type 16
ORGANISM Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
Alphapapillomavirus.
REFERENCE 1 (bases 1 to 451)
AUTHORS Watts,K.J., Thompson,C.H., Cossart,Y.E. and Rose,B.R.
TITLE Sequence variation and physical state of human papillomavirus type
16 cervical cancer isolates from Australia and New Caledonia
JOURNAL Int. J. Cancer 97 (6), 868-874 (2002)
PUBMED 11857370
REFERENCE 2 (bases 1 to 451)
AUTHORS Watts,K.J., Thompson,C.H., Cossart,Y.E. and Rose,B.R.
TITLE Direct Submission
JOURNAL Submitted (31-JUL-2001) Department of Infectious Diseases,
University of Sydney, Blackburn Building, D06, Off Western Avenue,
Camperdown, Sydney, New South Wales 2006, Australia
Location/Qualifiers
1..451
/organism="Human papillomavirus type 16"
/mol_type="genomic DNA"
/isolate="HPV16B6CC8"
/db_xref="taxon:333760"
<1..395
/gene="E6"
<1..395
/gene="E6"
/codon_start=3
/product="E6 protein"
/protein_id="AAL01355.1"
/db_xref="GI:15529591"
/translation="TTHNIIIECVYCKQQLRREVDPAFRDLCTVYRDGNPYAVCDK
CLKFYSKISSEYRYCVSVGTTLLEQYKNPCLDLLIRNCINQCKPLCPBEKQRHLDDKKO
RFINIRGRWTGRMCCSCRSSRTREITOL"
398..>451
/gene="E7"
398..>451
/codon_start=1
/product="E7 protein"
/protein_id="AAL01356.1"
/db_xref="GI:15529592"
/translation="MHGDTPTLHEMYMLDLPPE"

ORIGIN

Query Match 100.0%; Score 28; DB 10; Length 451;
Best Local Similarity 100.0%; Pred. No. 0.0025;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAACAGCAATACAAACCGTTGTGTG 28
|||||
204 GAACAGCAATACAAACCGTTGTGTG 231

Db

RESULT 18
AF404700

LOCUS AF04700 451 bp DNA linear VRL 09-OCT-2003
DEFINITION Human papillomavirus type 16 isolate HPV16ECC9 E6 protein (E6) and
E7 protein (E7) genes, partial cds.
ACCESSION AF04700
KEYWORDS AF04700.1 GI:15529593
SOURCE Human papillomavirus type 16
ORGANISM Human papillomavirus type 16
VIRUSES; dsDNA viruses, no RNA stage; Papillomaviridae;
Alphapapillomavirus.
REFERENCE 1 (bases 1 to 451)
AUTHORS Watts,K.J., Thompson,C.H., Cossart,Y.E. and Rose,B.R.
TITLE Sequence variation and physical state of human papillomavirus type
16 cervical cancer isolates from Australia and New Caledonia
JOURNAL Int. J. Cancer 97 (6), 868-874 (2002)
PUBMED 11857370
REFERENCE 2 (bases 1 to 451)
AUTHORS Watts,K.J., Thompson,C.H., Cossart,Y.E. and Rose,B.R.
TITLE Direct Submission
JOURNAL Submitted (31-JUL-2001) Department of Infectious Diseases,
University of Sydney, Blackburn Building, D06, Off Western Avenue,
Camperdown, Sydney, New South Wales 2006, Australia
Location/Qualifiers
1. .451
/organism="Human papillomavirus type 16"
/mol_type="genomic DNA"
/isolate="HPV16ECC9"
/db_xref="taxon:333760"
<1. .395
/gene="E6"
/gene="E7"
<1. .395
/gene="E6"
/gene="E7"
<1. .395
/product="E6 protein"
/codon_start=3
/protein_id="AA01357.1"
/db_xref="GI:15529594"
/translation="TTHDILLCVCYCKQLRREYDFAFRDICTIYRDGNPYAVCDK
CLMFYSKISEYRHYCYSVGTTLBOQYNPLCDLIRINCOKPLCPBEKORLDRKKQ
RPHNIGRWGTGRMCCSSRRTRRRTOL"
398. .>451
/gene="E7"
398. .>451
/gene="E7"
/codon_start=1
/product="E7 protein"
/protein_id="AA01358.1"
/db_xref="GI:15529595"
/translation="MHGDTPTLHRYMLDQPE"

ORIGIN
Query Match 100.0%; Score 28; DB 10; Length 451;
Best Local Similarity 100.0%; Pred. No. 0.0025;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAACAGCAATACAAACCGTGTGTG 28
|||||
Db 204 GAACAGCAATACAAACCGTGTGTG 231

RESULT 19
AF04701 451 bp DNA linear VRL 09-OCT-2003
LOCUS AF04701
DEFINITION Human papillomavirus type 16 isolate HPV16ECC10 E6 protein (E6)
and E7 protein (E7) genes, partial cds.
ACCESSION AF04701
KEYWORDS AF04701.1 GI:15529596
SOURCE Human papillomavirus type 16
ORGANISM Human papillomavirus type 16
VIRUSES; dsDNA viruses, no RNA stage; Papillomaviridae;
Alphapapillomavirus.
REFERENCE 1 (bases 1 to 451)
AUTHORS Watts,K.J., Thompson,C.H., Cossart,Y.E. and Rose,B.R.

TITLE Sequence variation and physical state of human papillomavirus type
16 cervical cancer isolates from Australia and New Caledonia
JOURNAL Int. J. Cancer 97 (6), 868-874 (2002)
PUBMED 11857370
REFERENCE 2 (bases 1 to 451)
AUTHORS Watts,K.J., Thompson,C.H., Cossart,Y.E. and Rose,B.R.
TITLE Direct Submission
JOURNAL Submitted (31-JUL-2001) Department of Infectious Diseases,
University of Sydney, Blackburn Building, D06, Off Western Avenue,
Camperdown, Sydney, New South Wales 2006, Australia
Location/Qualifiers
1. .451
/organism="Human papillomavirus type 16"
/mol_type="genomic DNA"
/isolate="HPV16ECC10"
/db_xref="taxon:333760"
<1. .395
/gene="E6"
/gene="E7"
<1. .395
/product="E6 protein"
/codon_start=3
/protein_id="AA01359.1"
/db_xref="GI:15529597"
/translation="TTHDILLCVCYCKQLRREYDFAFRDICTIYRDGNPYAVCDK
CLMFYSKISEYRHYCYSVGTTLBOQYNPLCDLIRINCOKPLCPBEKORLDRKKQ
RPHNIGRWGTGRMCCSSRRTRRRTOL"
398. .>451
/gene="E7"
398. .>451
/gene="E7"
/codon_start=1
/product="E7 protein"
/protein_id="AA01360.1"
/db_xref="GI:15529598"
/translation="MHGDTPTLHRYMLDQPE"

ORIGIN
Query Match 100.0%; Score 28; DB 10; Length 451;
Best Local Similarity 100.0%; Pred. No. 0.0025;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAACAGCAATACAAACCGTGTGTG 28
|||||
Db 204 GAACAGCAATACAAACCGTGTGTG 231

RESULT 20
AF04703 451 bp DNA linear VRL 09-OCT-2003
LOCUS AF04703
DEFINITION Human papillomavirus type 16 isolate HPV16ECC12 E6 protein (E6)
and E7 protein (E7) genes, partial cds.
ACCESSION AF04703
KEYWORDS AF04703.1 GI:15529602
SOURCE Human papillomavirus type 16
ORGANISM Human papillomavirus type 16
VIRUSES; dsDNA viruses, no RNA stage; Papillomaviridae;
Alphapapillomavirus.
REFERENCE 1 (bases 1 to 451)
AUTHORS Watts,K.J., Thompson,C.H., Cossart,Y.E. and Rose,B.R.
TITLE Sequence variation and physical state of human papillomavirus type
16 cervical cancer isolates from Australia and New Caledonia
JOURNAL Int. J. Cancer 97 (6), 868-874 (2002)
PUBMED 11857370
REFERENCE 2 (bases 1 to 451)
AUTHORS Watts,K.J., Thompson,C.H., Cossart,Y.E. and Rose,B.R.
TITLE Direct Submission
JOURNAL Submitted (31-JUL-2001) Department of Infectious Diseases,
University of Sydney, Blackburn Building, D06, Off Western Avenue,
Camperdown, Sydney, New South Wales 2006, Australia
Location/Qualifiers
1. .451

```

/organism="Human papillomavirus type 16"
/mol_type="genomic DNA"
/isolate="HPV16E6C12"
/db_xref="taxon:333760"
<1..395
/gene="E6"
<1..395
/gene="E6"
/codon_start=3
/product="E6 protein"
/protein_id="AAL01363.1"
/db_xref="GI:15529603"
/translation="TINHIILLCYCYCKOQLREVDFAFDLCTIVRDGNPYAVCDK
CIKYSKISEYRHYCSLYGTTLEOYNKPLCDLLIRINCQKPLCEBEKQRLHDKKQ
RFRNIRGRMTCRMCSCRSRTRRETQL"
398..>451
/gene="E7"
398..>451
/gene="E7"
/codon_start=1
/product="E7 protein"
/protein_id="AAL01364.1"
/db_xref="GI:15529604"
/translation="MHGDTPTLHEVMLDLQPE"

ORIGIN
Query Match 100.0%; Score 28; DB 10; Length 451;
Best Local Similarity 100.0%; Pred. No. 0.0025;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAACAGCAATACAAACCGTTGTGTG 28
Db 204 GAACAGCAATACAAACCGTTGTGTG 231

RESULT 21
ARI67393 456 bp DNA linear PAT 17-DEC-2001
LOCUS ARI67393
DEFINITION Sequence 29 from patent US 62875569.
ACCESSION ARI67393
VERSION ARI67393.1 GI:17903171
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 456)
AUTHORS Kippes,T.J. and Wu,Y.
TITLE Vaccines with enhanced intracellular processing
JOURNAL Patent: US 6287569-A 29 11-SEP-2001;
FEATURES
Location/Qualifiers
1..456
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 100.0%; Score 28; DB 2; Length 456;
Best Local Similarity 100.0%; Pred. No. 0.0025;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAACAGCAATACAAACCGTTGTGTG 28
Db 265 GAACAGCAATACAAACCGTTGTGTG 292

RESULT 22
ARI77943 456 bp DNA linear PAT 17-DEC-2001
LOCUS ARI77943
DEFINITION Sequence 6 from patent US 6313373.
ACCESSION ARI77943
VERSION ARI77943.1 GI:17920298
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.

```

```

Unclassified.
REFERENCE 1 (bases 1 to 456)
AUTHORS Eckert,R.L. and Crish,J.F.
TITLE Tissue specific promoters and transgenic mouse for the screening of
JOURNAL pharmaceuticals
Patent: US 6313373-A 6 06-NOV-2001;
FEATURES
Location/Qualifiers
1..456
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 100.0%; Score 28; DB 2; Length 456;
Best Local Similarity 100.0%; Pred. No. 0.0025;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAACAGCAATACAAACCGTTGTGTG 28
Db 265 GAACAGCAATACAAACCGTTGTGTG 292

RESULT 23
AF327851 456 bp DNA linear VRL 04-JAN-2001
LOCUS AF327851
DEFINITION Human papillomavirus type 16 early transforming protein E6 variant
ACCESSION AF327851
VERSION AF327851.1 GI:12025467
KEYWORDS Human papillomavirus type 16
SOURCE Human papillomavirus type 16
ORGANISM Human papillomavirus type 16
REFERENCE 1 (bases 1 to 456)
AUTHORS Ma,Z., Qian,D., Ma,J., Lin,R., Ming,W., Zhong,Z., Zhang,Q. and
Zhang,F.
TITLE Cloning and Sequencing of HPV16 E6 gene from Cervical Carcinoma
Biopsies in Xinjiang
JOURNAL Shengwu Huaxue Yu Shengwu Wuli Jinzhan (2001) In press
REFERENCE 2 (bases 1 to 456)
AUTHORS Ma,Z. and Zhang,F.
TITLE Direct Submision
JOURNAL Submitted (12-DEC-2000) Department of Biology, University of
Xinjiang, 14 Shengqi Road, Urumqi, Xinjiang 830046, P.R. China
FEATURES
Location/Qualifiers
1..456
/organism="Human papillomavirus type 16"
/mol_type="genomic DNA"
/mol_type="genomic DNA"
/strain="Xinjiang"
/specific_host="Homo sapiens"
/db_xref="taxon:333760"
/note="HPV16"
1..456
/gene="E6"
1..456
/gene="E6"
/codon_start=1
/product="early transforming protein E6 variant"
/protein_id="AAG45940.1"
/db_xref="GI:12025468"
/translation="MFQDPOERPRKLPOLCTELQTTIHDIILBCVYCKOQLRREVD
FAFRDLCTIVRDGNPYAVCDKCIKYSKISEYRHYCSLYGTTLEOYNKPLCDLLIR
CINQKPLCEBEKQRLHDKKQRFNIRGRMTCRMCSCRSRTRRETQL"

ORIGIN
Query Match 100.0%; Score 28; DB 10; Length 456;
Best Local Similarity 100.0%; Pred. No. 0.0025;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAACAGCAATACAAACCGTTGTGTG 28
Db 265 GAACAGCAATACAAACCGTTGTGTG 292

```

RESULT 24
LOCUS HPJ34107 456 bp DNA linear VRL 08-MAR-1996
DEFINITION Human papillomavirus type 16, isolate NM 4094, early transforming protein E6 (E6) gene, complete cds.
ACCESSION U34107
VERSION U34107.1 GI:1098719
KEYWORDS
SOURCE Human papillomavirus
ORGANISM Human papillomavirus
REFERENCE 1 (bases 1 to 456)
AUTHORS Yamada,T., Wheeler,C.M., Halpern,A.L., Stewart,A.C., Hildesheim,A. and Jensen,S.A.
TITLE Human papillomavirus type 16 variant lineages in United States populations characterized by nucleotide sequence analysis of the E6, E2, and L1 coding segments
J. Virol. 69 (12), 7743-7753 (1995)
JOURNAL 7494284
PUBMED 2 (bases 1 to 456)
REFERENCE Farmer,A.D.
AUTHORS Direct Submission
TITLE Submitted (16-AUG-1995) Andrew D. Farmer, Los Alamos National Laboratory, HIV Sequence Database, T-10, Mail Stop K710, Los Alamos, NM 87501, USA
LOCATION/Qualifiers
FEATURES
source
1..456
/organism="Human papillomavirus"
/mol_type="genomic DNA"
/isolate="NM 4094"
/db_xref="taxon:10566"
/note="collected by cervical swab at the University of New Mexico, New Mexico, United States."
1..456
/gene="E6"
1..456
/gene="E6"
/codon_start=1
/product="early transforming protein E6"
/protein_id="AA91654.1"
/db_xref="GI:1098720"
/translation="MFQDQERPRKLPHTLCTELQTTIHDIIEVCYCKOOLLREYVD FAPRDLCTVYRDGNPYAVCDKLPKSKSEYRYCYGYGTTLEOYVKKPLCDLLIR CINQKPLCEBERKHLDKQRFHNRGWTGRCHSCCSSRRRTQL"
ORIGIN
Query Match 100.0%; Score 28; DB 10; Length 456;
Best Local Similarity 100.0%; Pred. No. 0.0025;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GAACAGCAATACAAACCGTTGTGTG 28
Db 265 GAACAGCAATACAAACCGTTGTGTG 292
RESULT 25
LOCUS HPJ34108 456 bp DNA linear VRL 08-MAR-1996
DEFINITION Human papillomavirus type 16, isolate NM 9999, early transforming protein E6 (E6) gene, complete cds.
ACCESSION U34108
VERSION U34108.1 GI:1098721
KEYWORDS
SOURCE Human papillomavirus
ORGANISM Human papillomavirus
REFERENCE 1 (bases 1 to 456)
AUTHORS Yamada,T., Wheeler,C.M., Halpern,A.L., Stewart,A.C., Hildesheim,A. and Jensen,S.A.

TITLE Human papillomavirus type 16 variant lineages in United States populations characterized by nucleotide sequence analysis of the E6, E2, and L1 coding segments
J. Virol. 69 (12), 7743-7753 (1995)
JOURNAL 7494284
PUBMED 2 (bases 1 to 456)
REFERENCE Farmer,A.D.
AUTHORS Direct Submission
TITLE Submitted (16-AUG-1995) Andrew D. Farmer, Los Alamos National Laboratory, HIV Sequence Database, T-10, Mail Stop K710, Los Alamos, NM 87501, USA
LOCATION/Qualifiers
FEATURES
source
1..456
/organism="Human papillomavirus"
/mol_type="genomic DNA"
/isolate="NM 9999"
/db_xref="taxon:10566"
/note="collected by cervical swab at the University of New Mexico, New Mexico, United States."
1..456
/gene="E6"
1..456
/gene="E6"
/codon_start=1
/product="early transforming protein E6"
/protein_id="AA91655.1"
/db_xref="GI:1098722"
/translation="MFQDQERPRKLPHTLCTELQTTIHDIIEVCYCKOOLLREYVD FAPRDLCTVYRDGNPYAVCDKLPKSKSEYRYCYGYGTTLEOYVKKPLCDLLIR CINQKPLCEBERKHLDKQRFHNRGWTGRCHSCCSSRRRTQL"
ORIGIN
Query Match 100.0%; Score 28; DB 10; Length 456;
Best Local Similarity 100.0%; Pred. No. 0.0025;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GAACAGCAATACAAACCGTTGTGTG 28
Db 265 GAACAGCAATACAAACCGTTGTGTG 292
RESULT 26
LOCUS HPJ34109 456 bp DNA linear VRL 08-MAR-1996
DEFINITION Human papillomavirus type 16, isolate NM T197, early transforming protein E6 (E6) gene, complete cds.
ACCESSION U34109
VERSION U34109.1 GI:1098723
KEYWORDS
SOURCE Human papillomavirus
ORGANISM Human papillomavirus
REFERENCE 1 (bases 1 to 456)
AUTHORS Yamada,T., Wheeler,C.M., Halpern,A.L., Stewart,A.C., Hildesheim,A. and Jensen,S.A.
TITLE Human papillomavirus type 16 variant lineages in United States populations characterized by nucleotide sequence analysis of the E6, E2, and L1 coding segments
J. Virol. 69 (12), 7743-7753 (1995)
JOURNAL 7494284
PUBMED 2 (bases 1 to 456)
REFERENCE Farmer,A.D.
AUTHORS Direct Submission
TITLE Submitted (16-AUG-1995) Andrew D. Farmer, Los Alamos National Laboratory, HIV Sequence Database, T-10, Mail Stop K710, Los Alamos, NM 87501, USA
LOCATION/Qualifiers
FEATURES
source
1..456
/organism="Human papillomavirus"
/mol_type="genomic DNA"
/isolate="NM T197"
/db_xref="taxon:10566"

gene
/note="collected by cervical swab at the University of New Mexico, New Mexico, United States."
1. .456
/gene="E6"
1. .456
/gene="E6"
/codon_start=1
/product="early transforming protein E6"
/protein_id="AA91656.1"
/db_xref="GI:1098724"
/translation="MFQDPQRPKLPOLCTELOTTHDIIIECVCKOQLRREYD
FARPDCTIVRDGNPVAVCDKCKFKYSKISEYHYKCYSGTTLLEQYNNPLCDLLR
CINQKRLCEBEKQRHDKQRFNINGRWTGRCMSCRSRRTREROL"

ORIGIN
Query Match 100.0%; Score 28; DB 10; Length 456;
Best Local Similarity 100.0%; Pred. No. 0.0025;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 265 GAACAGCAATACAAACCGTTGTGTG 292

RESULT 27
HPJ34110 456 bp DNA linear VRL 08-MAR-1996
LOCUS Human papillomavirus type 16, isolate NM T446, early transforming
DEFINITION protein E6 (E6) gene, complete cds.
ACCESSION U34110
VERSION U34110.1 GI:1098725
KEYWORDS
SOURCE Human papillomavirus
ORGANISM Human papillomavirus
Virus; dsDNA viruses, no RNA stage; Papillomaviridae;
unclassified Papillomaviridae.
1 (bases 1 to 456)
Yamada,T., Wheeler,C.M., Halpern,A.L., Stewart,A.C., Hildesheim,A.
and Jenison,S.A.
Human papillomavirus type 16 variant lineages in United States
populations characterized by nucleotide sequence analysis of the
E6, L2, and L1 coding segments
J. Virol. 69 (12), 7743-7753 (1995)
7494284
2 (bases 1 to 456)
Farmer,A.D.
Direct Submision
Submitted (16-AUG-1995) Andrew D. Farmer, Los Alamos National
Laboratory, HIV Sequence Database, T-10, Mail Stop K710, Los
Alamos, NM 87501, USA
Location/Qualifiers
1. .456
/organism="Human papillomavirus"
/mol_type="genomic DNA"
/isolate="NM T446"
/db_xref="taxon:10566"
/note="collected by cervical swab at the University of New
Mexico, New Mexico, United States."
1. .456
/gene="E6"
1. .456
/gene="E6"
/codon_start=1
/product="early transforming protein E6"
/protein_id="AA91657.1"
/db_xref="GI:1098726"
/translation="MFQDPQRPKLPOLCTELOTTHDIIIECVCKOQLRREYD
FARPDCTIVRDGNPVAVCDKCKFKYSKISEYHYKCYSGTTLLEQYNNPLCDLLR
CINQKRLCEBEKQRHDKQRFNINGRWTGRCMSCRSRRTREROL"

ORIGIN
Query Match 100.0%; Score 28; DB 10; Length 456;
Best Local Similarity 100.0%; Pred. No. 0.0025;

Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAACAGCAATACAAACCGTTGTGTG 28
Db 265 GAACAGCAATACAAACCGTTGTGTG 292

RESULT 28
HPJ34111 456 bp DNA linear VRL 08-MAR-1996
LOCUS Human papillomavirus type 16, isolate NM T455, early transforming
DEFINITION protein E6 (E6) gene, complete cds.
ACCESSION U34111
VERSION U34111.1 GI:1098727
KEYWORDS
SOURCE Human papillomavirus
ORGANISM Human papillomavirus
Virus; dsDNA viruses, no RNA stage; Papillomaviridae;
unclassified Papillomaviridae.
1 (bases 1 to 456)
Yamada,T., Wheeler,C.M., Halpern,A.L., Stewart,A.C., Hildesheim,A.
and Jenison,S.A.
Human papillomavirus type 16 variant lineages in United States
populations characterized by nucleotide sequence analysis of the
E6, L2, and L1 coding segments
J. Virol. 69 (12), 7743-7753 (1995)
7494284
2 (bases 1 to 456)
Farmer,A.D.
Direct Submision
Submitted (16-AUG-1995) Andrew D. Farmer, Los Alamos National
Laboratory, HIV Sequence Database, T-10, Mail Stop K710, Los
Alamos, NM 87501, USA
Location/Qualifiers
1. .456
/organism="Human papillomavirus"
/mol_type="genomic DNA"
/isolate="NM T455"
/db_xref="taxon:10566"
/note="collected by cervical swab at the University of New
Mexico, New Mexico, United States."
1. .456
/gene="E6"
1. .456
/gene="E6"
/codon_start=1
/product="early transforming protein E6"
/protein_id="AA91658.1"
/db_xref="GI:1098728"
/translation="MFQDPQRPKLPOLCTELOTTHDIIIECVCKOQLRREYD
FARPDCTIVRDGNPVAVCDKCKFKYSKISEYHYKCYSGTTLLEQYNNPLCDLLR
CINQKRLCEBEKQRHDKQRFNINGRWTGRCMSCRSRRTREROL"

ORIGIN
Query Match 100.0%; Score 28; DB 10; Length 456;
Best Local Similarity 100.0%; Pred. No. 0.0025;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAACAGCAATACAAACCGTTGTGTG 28
Db 265 GAACAGCAATACAAACCGTTGTGTG 292

RESULT 29
HPJ34112 456 bp DNA linear VRL 08-MAR-1996
LOCUS Human papillomavirus type 16, isolate NM T529, early transforming
DEFINITION protein E6 (E6) gene, complete cds.
ACCESSION U34112
VERSION U34112.1 GI:1098729
KEYWORDS
SOURCE Human papillomavirus
ORGANISM Human papillomavirus

```

REFERENCE
AUTHORS      Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
              unclassified Papillomaviridae.
TITLE        1 (bases 1 to 456)
              Yamada,T., Wheeler,C.M., Halpern,A.L., Stewart,A.C., Hildesheim,A.
              and Jenson,S.A.
              Human papillomavirus type 16 variant lineages in United States
              populations characterized by nucleotide sequence analysis of the
              E6, E2, and L1 coding segments
JOURNAL      J. Virol. 69 (12), 7743-7753 (1995)
PUBMED      7494284
REFERENCE    2 (bases 1 to 456)
AUTHORS      Farmer,A.D.
              Direct Submission
              Submitted (16-AUG-1995) Andrew D. Farmer, Los Alamos National
              Laboratory, HIV Sequence Database, T-10, Mail Stop K710, Los
              Alamos, NM 87501, USA
FEATURES
  source
    1..456
      /organism="Human papillomavirus"
      /mol_type="genomic DNA"
      /isolate="NM T529"
      /db_xref="taxon:10566"
      /note="collected by cervical swab at the University of New
      Mexico, New Mexico, United States."
  gene
    1..456
      /gene="E6"
  CDS
    1..456
      /gene="E6"
      /codon_start=1
      /product="early transforming protein E6"
      /protein_id="AA91659.1"
      /db_xref="GI:1098730"
      /translation="MFODPQERPRKLPCLTELTQTTIHDIIECYCKQOQLRREYVD
      FARDCIIVRDGNPYAVCDKCLKFKYSKISERYCYSLYGTTLBOQYNKPLCDLLIR
      CINQKPLCEBKORHLDKORFHNIRGRTGCMSCSSRRRTQL"
ORIGIN
Query Match      100.0%; Score 28; DB 10; Length 456;
Best Local Similarity 100.0%; Pred. No. 0.0025;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY
  1 GAACAGCAATACAAACCGTTGTGTG 28
  |||||
  265 GAACAGCAATACAAACCGTTGTGTG 292
RESULT 30
LOCUS      HPU34113              456 bp      DNA      linear      VRL 08-MAR-1996
DEFINITION Human papillomavirus type 16, isolate OR 0198, early transforming
              protein E6 (E6) gene, complete cds.
VERSION    U34113
KEYWORDS   U34113.1 GI:1098731
SOURCE     Human papillomavirus
            Human papillomavirus
            Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
            unclassified Papillomaviridae.
REFERENCE  1 (bases 1 to 456)
AUTHORS    Yamada,T., Wheeler,C.M., Halpern,A.L., Stewart,A.C., Hildesheim,A.
            and Jenson,S.A.
            Human papillomavirus type 16 variant lineages in United States
            populations characterized by nucleotide sequence analysis of the
            E6, E2, and L1 coding segments
JOURNAL    J. Virol. 69 (12), 7743-7753 (1995)
PUBMED     7494284
REFERENCE  2 (bases 1 to 456)
AUTHORS    Farmer,A.D.
            Direct Submission
            Submitted (16-AUG-1995) Andrew D. Farmer, Los Alamos National
            Laboratory, HIV Sequence Database, T-10, Mail Stop K710, Los
            Alamos, NM 87501, USA
FEATURES
  source
    1..456
      /organism="Human papillomavirus"
      /mol_type="genomic DNA"
      /isolate="OR 1905"
      /db_xref="taxon:10566"
      /note="collected by cervical lavage at Kaiser Permanente,
      Portland, Oregon, United States."
  gene
    1..456
      /gene="E6"
  CDS
    1..456
      /gene="E6"
      /codon_start=1
      /product="early transforming protein E6"
      /protein_id="AA91661.1"
      /db_xref="GI:1098734"
      /translation="MFODPQERPRKLPCLTELTQTTIHDIIECYCKQOQLRREYVD
      FARDCIIVRDGNPYAVCDKCLKFKYSKISERYCYSLYGTTLBOQYNKPLCDLLIR
      CINQKPLCEBKORHLDKORFHNIRGRTGCMSCSSRRRTQL"

```

```

source
  1..456
    /organism="Human papillomavirus"
    /mol_type="genomic DNA"
    /isolate="OR 0198"
    /db_xref="taxon:10566"
    /note="collected by cervical lavage at Kaiser Permanente,
    Portland, Oregon, United States."
  gene
    1..456
      /gene="E6"
  CDS
    1..456
      /gene="E6"
      /codon_start=1
      /product="early transforming protein E6"
      /protein_id="AA91660.1"
      /db_xref="GI:1098732"
      /translation="MFODPQERPRKLPCLTELTQTTIHDIIECYCKQOQLRREYVD
      FARDCIIVRDGNPYAVCDKCLKFKYSKISERYCYSLYGTTLBOQYNKPLCDLLIR
      CINQKPLCEBKORHLDKORFHNIRGRTGCMSCSSRRRTQL"
ORIGIN
Query Match      100.0%; Score 28; DB 10; Length 456;
Best Local Similarity 100.0%; Pred. No. 0.0025;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY
  1 GAACAGCAATACAAACCGTTGTGTG 28
  |||||
  265 GAACAGCAATACAAACCGTTGTGTG 292
RESULT 31
LOCUS      HPU34114              456 bp      DNA      linear      VRL 08-MAR-1996
DEFINITION Human papillomavirus type 16, isolate OR 1905, early transforming
              protein E6 (E6) gene, complete cds.
VERSION    U34114
KEYWORDS   U34114.1 GI:1098733
SOURCE     Human papillomavirus
            Human papillomavirus
            Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
            unclassified Papillomaviridae.
REFERENCE  1 (bases 1 to 456)
AUTHORS    Yamada,T., Wheeler,C.M., Halpern,A.L., Stewart,A.C., Hildesheim,A.
            and Jenson,S.A.
            Human papillomavirus type 16 variant lineages in United States
            populations characterized by nucleotide sequence analysis of the
            E6, E2, and L1 coding segments
JOURNAL    J. Virol. 69 (12), 7743-7753 (1995)
PUBMED     7494284
REFERENCE  2 (bases 1 to 456)
AUTHORS    Farmer,A.D.
            Direct Submission
            Submitted (16-AUG-1995) Andrew D. Farmer, Los Alamos National
            Laboratory, HIV Sequence Database, T-10, Mail Stop K710, Los
            Alamos, NM 87501, USA
FEATURES
  source
    1..456
      /organism="Human papillomavirus"
      /mol_type="genomic DNA"
      /isolate="OR 1905"
      /db_xref="taxon:10566"
      /note="collected by cervical lavage at Kaiser Permanente,
      Portland, Oregon, United States."
  gene
    1..456
      /gene="E6"
  CDS
    1..456
      /gene="E6"
      /codon_start=1
      /product="early transforming protein E6"
      /protein_id="AA91661.1"
      /db_xref="GI:1098734"
      /translation="MFODPQERPRKLPCLTELTQTTIHDIIECYCKQOQLRREYVD
      FARDCIIVRDGNPYAVCDKCLKFKYSKISERYCYSLYGTTLBOQYNKPLCDLLIR
      CINQKPLCEBKORHLDKORFHNIRGRTGCMSCSSRRRTQL"

```


ORIGIN	CINORKPLCEBKQRHLDKQRFNNIKRGRTGRCMSCRSSRTRERQL"
Query Match	100.0%; Score 28; DB 10; Length 456;
Best Local Similarity	100.0%; Pred. No. 0.0025;
Matches	28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy	1 GAACGCAATACAAACAACCCTGTGTG 28
Db	265 GAACGCAATACAAACAACCCTGTGTG 292
RESULT 32	
HPU34115	456 bp DNA linear VRL 08-MAR-1996
LOCUS	Human papillomavirus type 16, isolate OR 2087, early transforming
DEFINITION	Human B6 (B6) gene, complete cds.
ACCESSION	U34115
VERSION	U34115.1 GI:1098735
KEYWORDS	
SOURCE	Human papillomavirus
ORGANISM	Human papillomavirus
REFERENCE	Viruses: dsDNA viruses, no RNA stage; Papillomaviridae; unclassified Papillomaviridae. 1 (bases 1 to 456)
AUTHORS	Yanada,T., Wheeler,C.M., Halpern,A.L., Stewart,A.C., Hildeshelm,A. and Jenison,S.A.
TITLE	Human papillomavirus type 16 variant lineages in United States populations characterized by nucleotide sequence analysis of the E6, U2, and L1 coding segments
JOURNAL	J. Virol. 69 (12), 7743-7753 (1995)
PUBMED	7494284
REFERENCE	2 (bases 1 to 456) Farmer,A.D.
AUTHORS	Direct Submission
TITLE	Submitted (15-AUG-1995) Andrew D. Farmer, Los Alamos National Laboratory, HIV Sequence Database, T-10, Mail Stop K710, Los Alamos, NM 87501, USA
JOURNAL	
FEATURES	Location/Qualifiers
source	1..456 /organism="Human papillomavirus" /mol_type="genomic DNA" /isolate="OR 2087" /db_xref="taxon:10566" /note="collected by cervical lavage at Kaiser Permanente, Portland, Oregon, United States." 1..456 /gene="E6" 1..456 /gene="E6" /codon_start=1 /product="early transforming protein E6" /protein_id="AAA91662.1" /db_xref="GI:1098736" translation="MFQDPERPKRLPOLCTELQTTHIILLCVCYCKQQLRRVYD FAFDLIVRDGNPAVCDCLEFKYSKISRYHCYSLYGTTLEQQYNKEPCDLLIR CINCPKLPCEBKQRHLDKQRFNNIKRGRTGRCMSCRSSRTRERQL"
gene	
CDS	
ORIGIN	
Query Match	100.0%; Score 28; DB 10; Length 456;
Best Local Similarity	100.0%; Pred. No. 0.0025;
Matches	28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy	1 GAACGCAATACAAACAACCCTGTGTG 28
Db	265 GAACGCAATACAAACAACCCTGTGTG 292
RESULT 33	
HPU34116	456 bp DNA linear VRL 08-MAR-1996
LOCUS	Human papillomavirus type 16, isolate OR 3136, early transforming
DEFINITION	protein E6 (E6) gene, complete cds.

ACCESSION	U34116	GI:1098737	
VERSION	U34116.1		
KEYWORDS			
SOURCE	Human papillomavirus		
ORGANISM	Human papillomavirus		
REFERENCE	1 (bases 1 to 456)		
AUTHORS	Yamada,T., Wheeler,C.M., Halpern,A.L., Stewart,A.C., Hildesheim,A. and Jenson,S.A.		
TITLE	Human papillomavirus type 16 variant lineages in United States populations characterized by nucleotide sequence analysis of the E6, E2, and L1 coding segments		
JOURNAL	J. Virol. 69 (12), 7743-7753 (1995)		
PUBMED	7494284		
REFERENCE	2 (bases 1 to 456)		
AUTHORS	Farmer,A.D.		
TITLE	Direct Submission		
JOURNAL	Submitted (16-AUG-1995) Andrew D. Farmer, Los Alamos National Laboratory, HIV Sequence Database, T-10, Mail Stop K710, Los Alamos, NM 87501, USA		
FEATURES	Location/Qualifiers		
source	1..456		
	/organism="Human papillomavirus"		
	/mol_type="genomic DNA"		
	/isolate="OR 3136"		
	/db_xref="taxon:10566"		
	/note="collected by cervical lavage at Kaiser Permanente, Portland, Oregon, United States."		
gene	1..456		
	/gene="E6"		
CDS	1..456		
	/gene="E6"		
	/codon_start=1		
	/product="early transforming protein E6"		
	/protein_id="AA81663.1"		
	/db_xref="GI:1098738"		
	/translation="MFQDQPERPRKPLPHLCTELQTTIHDIILCYCKQQLRREYVDFAFDLCTIVRDGNPVAVCDDKCPKFKYSIKSYCYSGYGTILEQYNPRLCDLIR CINQKRLCPBEKQRIHDKKQRFHNIRGMYTGRCMSCRSSRTRERQTL"		
ORIGIN			
Query Match	100.0%	Score 28;	DB 10; Length 456;
Best Local Similarity	100.0%;	Pred. No. 0.0025;	
Matches 28; Conservative	0;	Mismatches 0;	Indels 0; Gaps 0;
QY	1 GAACAGCAATTCACAAACCGTTGTGCG 28		
DB	265 GAACAGCAATTCACAAACCGTTGTGCG 292		
RESULT 34			
HPV34117	456 bp	DNA	linear
LOCUS	Human papillomavirus type 16, isolate OR 3473, early transforming protein E6 (E6) gene, complete cds.		
DEFINITION	U34117		
ACCESSION	U34117.1	GI:1098739	
VERSION			
KEYWORDS			
SOURCE	Human papillomavirus		
ORGANISM	Human papillomavirus		
REFERENCE	1 (bases 1 to 456)		
AUTHORS	Yamada,T., Wheeler,C.M., Halpern,A.L., Stewart,A.C., Hildesheim,A. and Jenson,S.A.		
TITLE	Human papillomavirus type 16 variant lineages in United States populations characterized by nucleotide sequence analysis of the E6, E2, and L1 coding segments		
JOURNAL	J. Virol. 69 (12), 7743-7753 (1995)		
PUBMED	7494284		
REFERENCE	2 (bases 1 to 456)		
AUTHORS	Farmer,A.D.		

TITLE Direct Submission
JOURNAL Submitted (16-AUG-1995) Andrew D. Farmer, Los Alamos National Laboratory, HIV Sequence Database, T-10, Mail Stop K710, Los Alamos, NM 87501, USA

FEATURES
source
1. .456
/organism="Human papillomavirus"
/mol_type="genomic DNA"
/isolate="OR 3473"
/db_xref="taxon:10566"
/note="collected by cervical lavage at Kaiser Permanente, Portland, Oregon, United States."

gene
1. .456
/gene="E6"

CDS
1. .456
/gene="E6"
/codon_start=1
/product="early transforming protein E6"
/protein_id="AA91664.1"
/db_xref="GI:1098740"
/translation="MFQDPQERPKLPDLCTELQTTIHDIILSCYCKQQLRREYDFAFRDLCIVRDGNPVAVCDCIKCFYSKISEYRYCYSLYGTLLBOQYNPLCDLLIR CINQKPLCEBKQRHLDKQRFHNIRGWTGCMSCCRSSRTRETOL"

ORIGIN
Query Match 100.0%; Score 28; DB 10; Length 456;
Best Local Similarity 100.0%; Pred. No. 0.0025;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAACAGCAATACAAACCGTTGTGTG 28
|||||
265 GAACAGCAATACAAACCGTTGTGTG 292

Db 265 GAACAGCAATACAAACCGTTGTGTG 292

RESULT 35
HPV34118 456 bp DNA linear VRL 08-MAR-1996
LOCUS Human papillomavirus type 16, isolate OR 3759, early transforming
DEFINITION protein E6 (E6) gene, complete cds.
U34118.1 GI:1098741

ACCESSION
U34118.1 GI:1098741

KEYWORDS
Human papillomavirus

SOURCE
Human papillomavirus

ORGANISM
Human papillomavirus
unclassified Papillomaviridae;
viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
unclassified Papillomaviridae.

REFERENCE
1 (bases 1 to 456)
Yamada,T., Wheeler,C.M., Halpern,A.L., Stewart,A.C., Hildesheim,A.
and Jensen,S.A.
Human papillomavirus type 16 variant lineages in United States
populations characterized by nucleotide sequence analysis of the
E6, E2, and L1 coding segments
J. Virol. 69 (12), 7743-7753 (1995)

TITLE
Human papillomavirus type 16 variant lineages in United States
populations characterized by nucleotide sequence analysis of the
E6, E2, and L1 coding segments

JOURNAL
J. Virol. 69 (12), 7743-7753 (1995)

REFERENCE
2 (bases 1 to 456)
Farmer,A.D.
Direct Submission
Submitted (16-AUG-1995) Andrew D. Farmer, Los Alamos National
Laboratory, HIV Sequence Database, T-10, Mail Stop K710, Los
Alamos, NM 87501, USA

FEATURES
source
1. .456
/organism="Human papillomavirus"
/mol_type="genomic DNA"
/isolate="OR 3759"
/db_xref="taxon:10566"
/note="collected by cervical lavage at Kaiser Permanente,
Portland, Oregon, United States."

gene
1. .456
/gene="E6"

CDS
1. .456
/gene="E6"
/codon_start=1

/product="early transforming protein E6"
/protein_id="AA91665.1"
/db_xref="GI:1098742"
/translation="MFQDPQERPKLPDLCTELQTTIHDIILSCYCKQQLRREYDFAFRDLCIVRDGNPVAVCDCIKCFYSKISEYRYCYSLYGTLLBOQYNPLCDLLIR CINQKPLCEBKQRHLDKQRFHNIRGWTGCMSCCRSSRTRETOL"

ORIGIN
Query Match 100.0%; Score 28; DB 10; Length 456;
Best Local Similarity 100.0%; Pred. No. 0.0025;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAACAGCAATACAAACCGTTGTGTG 28
|||||
265 GAACAGCAATACAAACCGTTGTGTG 292

Db 265 GAACAGCAATACAAACCGTTGTGTG 292

RESULT 36
HPV34119 456 bp DNA linear VRL 08-MAR-1996
LOCUS Human papillomavirus type 16, isolate OR 4541, early transforming
DEFINITION protein E6 (E6) gene, complete cds.
U34119.1 GI:1098743

ACCESSION
U34119.1 GI:1098743

KEYWORDS
Human papillomavirus

SOURCE
Human papillomavirus

ORGANISM
Human papillomavirus
unclassified Papillomaviridae;
viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
unclassified Papillomaviridae.

REFERENCE
1 (bases 1 to 456)
Yamada,T., Wheeler,C.M., Halpern,A.L., Stewart,A.C., Hildesheim,A.
and Jensen,S.A.
Human papillomavirus type 16 variant lineages in United States
populations characterized by nucleotide sequence analysis of the
E6, E2, and L1 coding segments
J. Virol. 69 (12), 7743-7753 (1995)

TITLE
Human papillomavirus type 16 variant lineages in United States
populations characterized by nucleotide sequence analysis of the
E6, E2, and L1 coding segments

JOURNAL
J. Virol. 69 (12), 7743-7753 (1995)

REFERENCE
2 (bases 1 to 456)
Farmer,A.D.
Direct Submission
Submitted (16-AUG-1995) Andrew D. Farmer, Los Alamos National
Laboratory, HIV Sequence Database, T-10, Mail Stop K710, Los
Alamos, NM 87501, USA

FEATURES
source
1. .456
/organism="Human papillomavirus"
/mol_type="genomic DNA"
/isolate="OR 4541"
/db_xref="taxon:10566"
/note="collected by cervical lavage at Kaiser Permanente,
Portland, Oregon, United States."

gene
1. .456
/gene="E6"

CDS
1. .456
/gene="E6"
/codon_start=1
/product="early transforming protein E6"
/protein_id="AA91666.1"
/db_xref="GI:1098744"
/translation="MFQDPQERPKLPDLCTELQTTIHDIILSCYCKQQLRREYDFAFRDLCIVRDGNPVAVCDCIKCFYSKISEYRYCYSLYGTLLBOQYNPLCDLLIR CINQKPLCEBKQRHLDKQRFHNIRGWTGCMSCCRSSRTRETOL"

ORIGIN
Query Match 100.0%; Score 28; DB 10; Length 456;
Best Local Similarity 100.0%; Pred. No. 0.0025;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAACAGCAATACAAACCGTTGTGTG 28
|||||
265 GAACAGCAATACAAACCGTTGTGTG 292

Db 265 GAACAGCAATACAAACCGTTGTGTG 292

RESULT 37
HPU34120 456 bp DNA linear VRL 08-MAR-1996
LOCUS Human papillomavirus type 16, isolate OR 4716, early transforming
DEFINITION protein E6 (E6) gene, complete cds.
ACCESSION U34120
VERSION U34120.1 GI:1098745
KEYWORDS
SOURCE Human papillomavirus
ORGANISM Human papillomavirus
VIRUSES; dsDNA viruses, no RNA stage; Papillomaviridae;
REFERENCE 1 (bases 1 to 456)
AUTHORS Yamada,T., Wheeler,C.M., Halpern,A.L., Stewart,A.C., Hildesheim,A.
and Jenison,S.A.
TITLE Human papillomavirus type 16 variant lineages in United States
populations characterized by nucleotide sequence analysis of the
E6, I2, and L1 coding segments
JOURNAL J. Virol. 69 (12), 7743-7753 (1995)
PUBMED 7494284
REFERENCE 2 (bases 1 to 456)
AUTHORS Farmer,A.D.
TITLE Direct Submission
SUBMITTED (16-AUG-1995) Andrew D. Farmer, Los Alamos National
Laboratory, HIV Sequence Database, T-10, Mail Stop K710, Los
Alamos, NM 87501, USA
LOCATION/Qualifiers
1. 456
/organism="Human papillomavirus"
/mol_type="genomic DNA"
/isolate="OR 4716"
/db_xref="taxon:10566"
/note="collected by cervical lavage at Kaiser Permanente,
Portland, Oregon, United States."
1. 456
/gene="E6"
1. 456
/gene="E6"
/codon_start=1
/product="early transforming protein E6"
/protein_id="AA91667.1"
/db_xref="GI:1098746"
/translation="MPQDPOBRPKLPOLCTELTTHIILIECVCKQOLREYVD
FARPLDCTIVRDGNPYAVCDKCLKFKYSIRYKCYSLYGTTEEGQYKRLCDLLR
CINQKPLCEBKORHLDKORFHNIRMTGRMCSRSSRTRETL"

ORIGIN
Query Match 100.0%; Score 28; DB 10; Length 456;
Best Local Similarity 100.0%; Pred. No. 0.0025;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 1 GAACGACGATACCAACCAACCGTTGTG 28
|||||
Db 265 GAACGACGATACCAACCAACCGTTGTG 292
|||||

RESULT 38
HPU34121 456 bp DNA linear VRL 08-MAR-1996
LOCUS Human papillomavirus type 16, isolate OR 4724, early transforming
DEFINITION protein E6 (E6) gene, complete cds.
ACCESSION U34121
VERSION U34121.1 GI:1098747
KEYWORDS
SOURCE Human papillomavirus
ORGANISM Human papillomavirus
VIRUSES; dsDNA viruses, no RNA stage; Papillomaviridae;
REFERENCE 1 (bases 1 to 456)
AUTHORS Yamada,T., Wheeler,C.M., Halpern,A.L., Stewart,A.C., Hildesheim,A.
and Jenison,S.A.
TITLE Human papillomavirus type 16 variant lineages in United States
populations characterized by nucleotide sequence analysis of the

JOURNAL J. Virol. 69 (12), 7743-7753 (1995)
PUBMED 7494284
REFERENCE 2 (bases 1 to 456)
AUTHORS Farmer,A.D.
TITLE Direct Submission
SUBMITTED (16-AUG-1995) Andrew D. Farmer, Los Alamos National
Laboratory, HIV Sequence Database, T-10, Mail Stop K710, Los
Alamos, NM 87501, USA
LOCATION/Qualifiers
1. 456
/organism="Human papillomavirus"
/mol_type="genomic DNA"
/isolate="OR 4724"
/db_xref="taxon:10566"
/note="collected by cervical lavage at Kaiser Permanente,
Portland, Oregon, United States."
1. 456
/gene="E6"
1. 456
/gene="E6"
/codon_start=1
/product="early transforming protein E6"
/protein_id="AA91668.1"
/db_xref="GI:1098748"
/translation="MPQDPOBRPKLPOLCTELTTHIILIECVCKQOLREYVD
FARPLDCTIVRDGNPYAVCDKCLKFKYSIRYKCYSLYGTTEEGQYKRLCDLLR
CINQKPLCEBKORHLDKORFHNIRMTGRMCSRSSRTRETL"

ORIGIN
Query Match 100.0%; Score 28; DB 10; Length 456;
Best Local Similarity 100.0%; Pred. No. 0.0025;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 1 GAACGACGATACCAACCAACCGTTGTG 28
|||||
Db 265 GAACGACGATACCAACCAACCGTTGTG 292
|||||

RESULT 39
HPU34122 456 bp DNA linear VRL 08-MAR-1996
LOCUS Human papillomavirus type 16, isolate OR 4997, early transforming
DEFINITION protein E6 (E6) gene, complete cds.
ACCESSION U34122
VERSION U34122.1 GI:1098749
KEYWORDS
SOURCE Human papillomavirus
ORGANISM Human papillomavirus
VIRUSES; dsDNA viruses, no RNA stage; Papillomaviridae;
REFERENCE 1 (bases 1 to 456)
AUTHORS Yamada,T., Wheeler,C.M., Halpern,A.L., Stewart,A.C., Hildesheim,A.
and Jenison,S.A.
TITLE Human papillomavirus type 16 variant lineages in United States
populations characterized by nucleotide sequence analysis of the
E6, I2, and L1 coding segments
JOURNAL J. Virol. 69 (12), 7743-7753 (1995)
PUBMED 7494284
REFERENCE 2 (bases 1 to 456)
AUTHORS Farmer,A.D.
TITLE Direct Submission
SUBMITTED (16-AUG-1995) Andrew D. Farmer, Los Alamos National
Laboratory, HIV Sequence Database, T-10, Mail Stop K710, Los
Alamos, NM 87501, USA
LOCATION/Qualifiers
1. 456
/organism="Human papillomavirus"
/mol_type="genomic DNA"
/isolate="OR 4997"
/db_xref="taxon:10566"
/note="collected by cervical lavage at Kaiser Permanente,
Portland, Oregon, United States."

gene 1. .456
/gene="B6"
CDS 1. .456
/gene="B6"
/codon_start=1
/product="early transforming protein B6"
/protein_id="AA91669.1"
/db_xref="GI:1098750"
/translation="MFODPQERPRKLPOLCTELQTTIHILLCVCYCKOOLRREYVD
FARDCIVYRDGNPYAVCDKCKIFYSKISRYHVCYSYGTLLBQYINPPLCDLLIR
CINCKPPLCEBKHLDKORPHNIRGWTGCMSCCRSRRTRRETOL"

ORIGIN

Query Match 100.0%; Score 28; DB 10; Length 456;
Best Local Similarity 100.0%; Pred. No. 0.0025;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAACAGCAATACAAACCGTTGTGTG 28
|||||
265 GAACAGCAATACAAACCGTTGTGTG 292

Db

RESULT 40
HPJ34123 456 bp DNA linear VRL 08-MAR-1996
LOCUS Human papillomavirus type 16, isolate OR 5110, early transforming
DEFINITION protein B6 (B6) gene, complete cds.
ACCESSION U34123
VERSION U34123.1 GI:1098751
KEYWORDS
SOURCE Human papillomavirus
ORGANISM Human papillomavirus
Virusess; dsDNA viruses, no RNA stage; Papillomaviridae;
unclassified Papillomaviridae.
REFERENCE 1 (bases 1 to 456)
Yamada,T., Wheeler,C.M., Halpern,A.L., Stewart,A.C., Hildesheim,A.
and Jenison,S.A.
Human papillomavirus type 16 variant lineages in United States
populations characterized by nucleotide sequence analysis of the
E6, E2, and L1 coding segments
J. Virol. 69 (12), 7743-7753 (1995)
7494284
2 (bases 1 to 456)
Farmer,A.D.
Direct Submission
Submitted (16-AUG-1995) Andrew D. Farmer, Los Alamos National
Laboratory, HIV Sequence Database, T-10, Mail Stop K710, Los
Alamos, NM 87501, USA
Location/Qualifiers
1. .456
/organism="Human papillomavirus"
/mol_type="genomic DNA"
/isolate="OR 5110"
/db_xref="taxon:10566"
/note="collected by cervical lavage at Kaiser Permanente,
Portland, Oregon, United States."
1. .456
/gene="B6"
1. .456
/gene="B6"
/codon_start=1
/product="early transforming protein B6"
/protein_id="AA91670.1"
/db_xref="GI:1098752"
/translation="MFODPQERPRKLPOLCTELQTTIHILLCVCYCKOOLRREYVD
FARDCIVYRDGNPYAVCDKCKIFYSKISRYHVCYSYGTLLBQYINPPLCDLLIR
CINCKPPLCEBKHLDKORPHNIRGWTGCMSCCRSRRTRRETOL"

ORIGIN

Query Match 100.0%; Score 28; DB 10; Length 456;
Best Local Similarity 100.0%; Pred. No. 0.0025;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAACAGCAATACAAACCGTTGTGTG 28
|||||
265 GAACAGCAATACAAACCGTTGTGTG 292

Db

RESULT 41
HPJ34124 456 bp DNA linear VRL 08-MAR-1996
LOCUS Human papillomavirus type 16, isolate OR 5428, early transforming
DEFINITION protein B6 (B6) gene, complete cds.
ACCESSION U34124
VERSION U34124.1 GI:1098753
KEYWORDS
SOURCE Human papillomavirus
ORGANISM Human papillomavirus
Virusess; dsDNA viruses, no RNA stage; Papillomaviridae;
unclassified Papillomaviridae.
REFERENCE 1 (bases 1 to 456)
Yamada,T., Wheeler,C.M., Halpern,A.L., Stewart,A.C., Hildesheim,A.
and Jenison,S.A.
Human papillomavirus type 16 variant lineages in United States
populations characterized by nucleotide sequence analysis of the
E6, E2, and L1 coding segments
J. Virol. 69 (12), 7743-7753 (1995)
7494284
2 (bases 1 to 456)
Farmer,A.D.
Direct Submission
Submitted (16-AUG-1995) Andrew D. Farmer, Los Alamos National
Laboratory, HIV Sequence Database, T-10, Mail Stop K710, Los
Alamos, NM 87501, USA
Location/Qualifiers
1. .456
/organism="Human papillomavirus"
/mol_type="genomic DNA"
/isolate="OR 5428"
/db_xref="taxon:10566"
/note="collected by cervical lavage at Kaiser Permanente,
Portland, Oregon, United States."
1. .456
/gene="B6"
1. .456
/gene="B6"
/codon_start=1
/product="early transforming protein B6"
/protein_id="AA91671.1"
/db_xref="GI:1098754"
/translation="MFODPQERPRKLPOLCTELQTTIHILLCVCYCKOOLRREYVD
FARDCIVYRDGNPYAVCDKCKIFYSKISRYHVCYSYGTLLBQYINPPLCDLLIR
CINCKPPLCEBKHLDKORPHNIRGWTGCMSCCRSRRTRRETOL"

ORIGIN

Query Match 100.0%; Score 28; DB 10; Length 456;
Best Local Similarity 100.0%; Pred. No. 0.0025;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAACAGCAATACAAACCGTTGTGTG 28
|||||
265 GAACAGCAATACAAACCGTTGTGTG 292

Db

RESULT 42
HPJ34125 456 bp DNA linear VRL 08-MAR-1996
LOCUS Human papillomavirus type 16, isolate OR 6106, early transforming
DEFINITION protein B6 (B6) gene, complete cds.
ACCESSION U34125
VERSION U34125.1 GI:1098755
KEYWORDS
SOURCE Human papillomavirus
ORGANISM Human papillomavirus
Virusess; dsDNA viruses, no RNA stage; Papillomaviridae;
unclassified Papillomaviridae.

REFERENCE 1 (bases 1 to 456)
AUTHORS Yamada,T., Wheeler,C.M., Halpern,A.L., Stewart,A.C., Hildesheim,A.
and Jenison,S.A.
TITLE Human papillomavirus type 16 variant lineages in United States
populations characterized by nucleotide sequence analysis of the
E6, E2, and L1 coding segments
JOURNAL J. Virol. 69 (12), 7743-7753 (1995)
PUBMED 7494284
REFERENCE 2 (bases 1 to 456)
AUTHORS Farmer,A.D.
TITLE Direct Submission
JOURNAL Submitted (16-AUG-1995) Andrew D. Farmer, Los Alamos National
Laboratory, HIV Sequence Database, T-10, Mail Stop K710, Los
Alamos, NM 87501, USA
FEATURES
SOURCE Location/Qualifiers
1..456
/organism="Human papillomavirus"
/mol_type="genomic DNA"
/isolate="OR 6170"
/db_xref="taxon:10566"
/note="collected by cervical lavage at Kaiser Permanente,
Portland, Oregon, United States."
gene 1..456
/gene="E6"
1..456
/gene="E6"
/codon_start=1
/product="early transforming protein E6"
/protein_id="AA91672.1"
/db_xref="GI:1098756"
/translation="MFODPQRPRLPDLCTELQTTIHDIILCEVYCKQQLRREYD
FAPRDLCTIVRDGNPVAVCCKLKYKISBYRYCYVGTTLBOQYNPLCDLLR
CINCQKPLCEBKORHLDKORFHNIRGRTGRCMSCRSSRTRERQL"
ORIGIN
Query Match 100.0%; Score 28; DB 10; Length 456;
Best Local Similarity 100.0%; Pred. No. 0.0025;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GAACAGCAATPACAAACCGTGTGTG 28
|||||
Db 265 GAACAGCAATPACAAACCGTGTGTG 292
RESULT 43
HPJ34126 456 bp DNA linear VRL 08-MAR-1996
LOCUS Human papillomavirus type 16, isolate OR 6170, early transforming
DEFINITION protein E6 (E6) gene, complete cds.
ACCESSION U34126
VERSION U34126.1 GI:1098757
KEYWORDS Human papillomavirus
SOURCE Human papillomavirus
ORGANISM Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
unclassified Papillomaviridae.
REFERENCE 1 (bases 1 to 456)
AUTHORS Yamada,T., Wheeler,C.M., Halpern,A.L., Stewart,A.C., Hildesheim,A.
and Jenison,S.A.
TITLE Human papillomavirus type 16 variant lineages in United States
populations characterized by nucleotide sequence analysis of the
E6, E2, and L1 coding segments
JOURNAL J. Virol. 69 (12), 7743-7753 (1995)
PUBMED 7494284
REFERENCE 2 (bases 1 to 456)
AUTHORS Farmer,A.D.
TITLE Direct Submission
JOURNAL Submitted (16-AUG-1995) Andrew D. Farmer, Los Alamos National
Laboratory, HIV Sequence Database, T-10, Mail Stop K710, Los
Alamos, NM 87501, USA
FEATURES
SOURCE Location/Qualifiers
1..456
/organism="Human papillomavirus"
/mol_type="genomic DNA"
/isolate="OR 6170"
/db_xref="taxon:10566"
/note="collected by cervical lavage at Kaiser Permanente,
Portland, Oregon, United States."
gene 1..456
/gene="E6"
1..456
/gene="E6"
/codon_start=1
/product="early transforming protein E6"
/protein_id="AA91674.1"
/db_xref="GI:1098760"
/translation="MFODPQRPRLPDLCTELQTTIHDIILCEVYCKQQLRREYD
FAPRDLCTIVRDGNPVAVCCKLKYKISBYRYCYVGTTLBOQYNPLCDLLR
CINCQKPLCEBKORHLDKORFHNIRGRTGRCMSCRSSRTRERQL"
ORIGIN
Query Match 100.0%; Score 28; DB 10; Length 456;
Best Local Similarity 100.0%; Pred. No. 0.0025;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GAACAGCAATPACAAACCGTGTGTG 28
|||||
Db 265 GAACAGCAATPACAAACCGTGTGTG 292
RESULT 44
HPJ34127 456 bp DNA linear VRL 08-MAR-1996
LOCUS Human papillomavirus type 16, isolate OR 6311, early transforming
DEFINITION protein E6 (E6) gene, complete cds.
ACCESSION U34127
VERSION U34127.1 GI:1098759
KEYWORDS Human papillomavirus
SOURCE Human papillomavirus
ORGANISM Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
unclassified Papillomaviridae.
REFERENCE 1 (bases 1 to 456)
AUTHORS Yamada,T., Wheeler,C.M., Halpern,A.L., Stewart,A.C., Hildesheim,A.
and Jenison,S.A.
TITLE Human papillomavirus type 16 variant lineages in United States
populations characterized by nucleotide sequence analysis of the
E6, E2, and L1 coding segments
JOURNAL J. Virol. 69 (12), 7743-7753 (1995)
PUBMED 7494284
REFERENCE 2 (bases 1 to 456)
AUTHORS Farmer,A.D.
TITLE Direct Submission
JOURNAL Submitted (16-AUG-1995) Andrew D. Farmer, Los Alamos National
Laboratory, HIV Sequence Database, T-10, Mail Stop K710, Los
Alamos, NM 87501, USA
FEATURES
SOURCE Location/Qualifiers
1..456
/organism="Human papillomavirus"
/mol_type="genomic DNA"
/isolate="OR 6311"
/db_xref="taxon:10566"
/note="collected by cervical lavage at Kaiser Permanente,
Portland, Oregon, United States."
gene 1..456
/gene="E6"
1..456
/gene="E6"
/codon_start=1
/product="early transforming protein E6"
/protein_id="AA91674.1"
/db_xref="GI:1098760"
/translation="MFODPQRPRLPDLCTELQTTIHDIILCEVYCKQQLRREYD
FAPRDLCTIVRDGNPVAVCCKLKYKISBYRYCYVGTTLBOQYNPLCDLLR
CINCQKPLCEBKORHLDKORFHNIRGRTGRCMSCRSSRTRERQL"
ORIGIN

REFERENCE 1 (bases 1 to 456)
AUTHORS Yamada,T., Wheeler,C.M., Halpern,A.L., Stewart,A.C., Hildesheim,A.
and Jenison,S.A.
TITLE Human papillomavirus type 16 variant lineages in United States
populations characterized by nucleotide sequence analysis of the
E6, E2, and L1 coding segments
JOURNAL J. Virol. 69 (12), 7743-7753 (1995)
PUBMED 7494284
REFERENCE 2 (bases 1 to 456)
AUTHORS Farmer,A.D.
TITLE Direct Submission
JOURNAL Submitted (16-AUG-1995) Andrew D. Farmer, Los Alamos National
Laboratory, HIV Sequence Database, T-10, Mail Stop K710, Los
Alamos, NM 87501, USA
FEATURES
SOURCE Location/Qualifiers
1..456
/organism="Human papillomavirus"
/mol_type="genomic DNA"
/isolate="OR 6170"
/db_xref="taxon:10566"
/note="collected by cervical lavage at Kaiser Permanente,
Portland, Oregon, United States."
gene 1..456
/gene="E6"
1..456
/gene="E6"
/codon_start=1
/product="early transforming protein E6"
/protein_id="AA91673.1"
/db_xref="GI:1098758"
/translation="MFODPQRPRLPDLCTELQTTIHDIILCEVYCKQQLRREYD
FAPRDLCTIVRDGNPVAVCCKLKYKISBYRYCYVGTTLBOQYNPLCDLLR
CINCQKPLCEBKORHLDKORFHNIRGRTGRCMSCRSSRTRERQL"
ORIGIN
Query Match 100.0%; Score 28; DB 10; Length 456;
Best Local Similarity 100.0%; Pred. No. 0.0025;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GAACAGCAATPACAAACCGTGTGTG 28
|||||
Db 265 GAACAGCAATPACAAACCGTGTGTG 292
RESULT 44
HPJ34127 456 bp DNA linear VRL 08-MAR-1996
LOCUS Human papillomavirus type 16, isolate OR 6311, early transforming
DEFINITION protein E6 (E6) gene, complete cds.
ACCESSION U34127
VERSION U34127.1 GI:1098759
KEYWORDS Human papillomavirus
SOURCE Human papillomavirus
ORGANISM Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
unclassified Papillomaviridae.
REFERENCE 1 (bases 1 to 456)
AUTHORS Yamada,T., Wheeler,C.M., Halpern,A.L., Stewart,A.C., Hildesheim,A.
and Jenison,S.A.
TITLE Human papillomavirus type 16 variant lineages in United States
populations characterized by nucleotide sequence analysis of the
E6, E2, and L1 coding segments
JOURNAL J. Virol. 69 (12), 7743-7753 (1995)
PUBMED 7494284
REFERENCE 2 (bases 1 to 456)
AUTHORS Farmer,A.D.
TITLE Direct Submission
JOURNAL Submitted (16-AUG-1995) Andrew D. Farmer, Los Alamos National
Laboratory, HIV Sequence Database, T-10, Mail Stop K710, Los
Alamos, NM 87501, USA
FEATURES
SOURCE Location/Qualifiers
1..456
/organism="Human papillomavirus"
/mol_type="genomic DNA"
/isolate="OR 6311"
/db_xref="taxon:10566"
/note="collected by cervical lavage at Kaiser Permanente,
Portland, Oregon, United States."
gene 1..456
/gene="E6"
1..456
/gene="E6"
/codon_start=1
/product="early transforming protein E6"
/protein_id="AA91674.1"
/db_xref="GI:1098760"
/translation="MFODPQRPRLPDLCTELQTTIHDIILCEVYCKQQLRREYD
FAPRDLCTIVRDGNPVAVCCKLKYKISBYRYCYVGTTLBOQYNPLCDLLR
CINCQKPLCEBKORHLDKORFHNIRGRTGRCMSCRSSRTRERQL"
ORIGIN

Query Match 100.0%; Score 28; DB 10; Length 456;
 Best Local Similarity 100.0%; Pred. No. 0.0025;
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAACAGCAATACAAACCGTTGTGTG 28
 |||||
 Db 265 GAACAGCAATACAAACCGTTGTGTG 292

RESULT 45
 HPJ34128 456 bp DNA linear VRL 08-MAR-1996
 LOCUS Human papillomavirus type 16, isolate OR 7145, early transforming
 DEFINITION protein E6 (E6) gene, complete cds.
 ACCESSION U34128
 KEYWORDS
 SOURCE Human papillomavirus
 ORGANISM Human papillomavirus
 unclassified Papillomaviridae;

REFERENCE 1 (bases 1 to 456)
 Yamada,T., Wheeler,C.M., Halpern,A.L., Stewart,A.C., Hildesheim,A.
 and Jenison,S.A.
 Human papillomavirus type 16 variant lineages in United States
 populations characterized by nucleotide sequence analysis of the
 E6, L2, and L1 coding segments
 J. Virol. 69 (12), 7743-7753 (1995)

JOURNAL 7494284
 PUBMED 2 (bases 1 to 456)
 AUTHORS Farmer,A.D.
 TITLE Direct Submission
 JOURNAL Submitted (16-AUG-1995) Andrew D. Farmer, Los Alamos National
 Laboratory, HIV Sequence Database, T-10, Mail Stop K710, Los
 Alamos, NM 87501, USA

FEATURES
 source Location/Qualifiers
 1..456
 /organism="Human papillomavirus"
 /mol_type="genomic DNA"
 /isolate="OR 7145"
 /db_xref="taxon:10566"
 /note="collected by cervical lavage at Kaiser Permanente,
 Portland, Oregon, United States."
 1..456
 /gene="E6"
 1..456
 /gene="E6"
 /codon_start=1
 /product="early transforming protein E6"
 /protein_id="AA91675.1"
 /db_xref="GI:1098763"
 /translations="MFQDPQERPKLPDCTELQTTIHILIEVCYCKQQLRREYD
 FAFRDLCTIVRDGNPYAVCDKLFKFSKSEYRYCYSLYGTLLDQYNNPLCDLLIR
 CINQKPLCPBEKQRHLDKQRFHNIRGWTGRCMSCRSSRTRRETQL"

ORIGIN
 Query Match 100.0%; Score 28; DB 10; Length 456;
 Best Local Similarity 100.0%; Pred. No. 0.0025;
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAACAGCAATACAAACCGTTGTGTG 28
 |||||
 Db 265 GAACAGCAATACAAACCGTTGTGTG 292

RESULT 46
 HPJ34129 456 bp DNA linear VRL 08-MAR-1996
 LOCUS Human papillomavirus type 16, isolate OR 7574, early transforming
 DEFINITION protein E6 (E6) gene, complete cds.
 ACCESSION U34129
 KEYWORDS
 SOURCE Human papillomavirus
 ORGANISM Human papillomavirus
 unclassified Papillomaviridae;

REFERENCE 1 (bases 1 to 456)
 Yamada,T., Wheeler,C.M., Halpern,A.L., Stewart,A.C., Hildesheim,A.
 and Jenison,S.A.
 Human papillomavirus type 16 variant lineages in United States
 populations characterized by nucleotide sequence analysis of the
 E6, L2, and L1 coding segments
 J. Virol. 69 (12), 7743-7753 (1995)

JOURNAL 7494284
 PUBMED 2 (bases 1 to 456)
 AUTHORS Farmer,A.D.
 TITLE Direct Submission
 JOURNAL Submitted (16-AUG-1995) Andrew D. Farmer, Los Alamos National

KEYWORDS
 SOURCE Human papillomavirus
 ORGANISM Human papillomavirus
 unclassified Papillomaviridae;

REFERENCE 1 (bases 1 to 456)
 Yamada,T., Wheeler,C.M., Halpern,A.L., Stewart,A.C., Hildesheim,A.
 and Jenison,S.A.
 Human papillomavirus type 16 variant lineages in United States
 populations characterized by nucleotide sequence analysis of the
 E6, L2, and L1 coding segments
 J. Virol. 69 (12), 7743-7753 (1995)

JOURNAL 7494284
 PUBMED 2 (bases 1 to 456)
 AUTHORS Farmer,A.D.
 TITLE Direct Submission
 JOURNAL Submitted (16-AUG-1995) Andrew D. Farmer, Los Alamos National

FEATURES
 source Location/Qualifiers
 1..456
 /organism="Human papillomavirus"
 /mol_type="genomic DNA"
 /isolate="OR 7574"
 /db_xref="taxon:10566"
 /note="collected by cervical lavage at Kaiser Permanente,
 Portland, Oregon, United States."
 1..456
 /gene="E6"
 1..456
 /gene="E6"
 /codon_start=1
 /product="early transforming protein E6"
 /protein_id="AA91676.1"
 /db_xref="GI:1098764"
 /translations="MFQDPQERPKLPDCTELQTTIHILIEVCYCKQQLRREYD
 FAFRDLCTIVRDGNPYAVCDKLFKFSKSEYRYCYSLYGTLLDQYNNPLCDLLIR
 CINQKPLCPBEKQRHLDKQRFHNIRGWTGRCMSCRSSRTRRETQL"

ORIGIN
 Query Match 100.0%; Score 28; DB 10; Length 456;
 Best Local Similarity 100.0%; Pred. No. 0.0025;
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAACAGCAATACAAACCGTTGTGTG 28
 |||||
 Db 265 GAACAGCAATACAAACCGTTGTGTG 292

RESULT 47
 HPJ34130 456 bp DNA linear VRL 08-MAR-1996
 LOCUS Human papillomavirus type 16, isolate OR 7587, early transforming
 DEFINITION protein E6 (E6) gene, complete cds.
 ACCESSION U34130
 KEYWORDS
 SOURCE Human papillomavirus
 ORGANISM Human papillomavirus
 unclassified Papillomaviridae;

REFERENCE 1 (bases 1 to 456)
 Yamada,T., Wheeler,C.M., Halpern,A.L., Stewart,A.C., Hildesheim,A.
 and Jenison,S.A.
 Human papillomavirus type 16 variant lineages in United States
 populations characterized by nucleotide sequence analysis of the
 E6, L2, and L1 coding segments
 J. Virol. 69 (12), 7743-7753 (1995)

JOURNAL 7494284
 PUBMED 2 (bases 1 to 456)
 AUTHORS Farmer,A.D.
 TITLE Direct Submission
 JOURNAL Submitted (16-AUG-1995) Andrew D. Farmer, Los Alamos National

FEATURES
source
Laboratory, HIV Sequence Database, T-10, Mail Stop K710, Los Alamos, NM 87501, USA
location/Qualifiers
1..456
/organism="Human papillomavirus"
/mol_type="genomic DNA"
/isolate="OR 7587"
/db_xref="taxon:10566"
/note="collected by cervical lavage at Kaiser Permanente, Portland, Oregon, United States."
1..456
/gene="E6"
1..456
/gene="E6"
/codon_start=1
/product="early transforming protein E6"
/protein_id="AA91677.1"
/db_xref="GI:1098766"
/translation="MFQDPERPRKLPDLCTELQTTIHDIILCEVYCKQQLRREYD FAPRDLCTIVRDGNPYAVCDCKFKYSKISERYCYSLVGTLLQOYNKPLCDLLR CINQKRLCEBKQRHDKQRPHNIRGRTGRCMSCRSRRTRETOL"

ORIGIN
Query Match 100.0%; Score 28; DB 10; Length 456;
Best Local Similarity 100.0%; Pred. No. 0.0025;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAACAGCAATACAAACCGTTGTG 28
|||||
265 GAACAGCAATACAAACCGTTGTG 292

Db 265 GAACAGCAATACAAACCGTTGTG 292

RESULT 48
HPU34131 456 bp DNA linear VRL 08-MAR-1996
LOCUS Human papillomavirus type 16, isolate OR 7632, early transforming
DEFINITION protein E6 (E6) gene, complete cds.
VERSION U34131.1 GI:1098767
KEYWORDS Human papillomavirus
SOURCE Human papillomavirus
ORGANISM Human papillomavirus
unclassified Papillomaviridae;
REFERENCE 1 (bases 1 to 456)
AUTHORS Yamada,T., Wheeler,C.M., Halpern,A.L., Stewart,A.C., Hildesheim,A. and Jensen,S.A.
TITLE Human papillomavirus type 16 variant lineages in United States populations characterized by nucleotide sequence analysis of the E6, E2, and L1 coding segments
JOURNAL U. Virol. 69 (12), 7743-7753 (1995)
PUBMED 7494284
REFERENCE 2 (bases 1 to 456)
AUTHORS Farmer,A.D.
TITLE Direct Submission
JOURNAL Submitted (16-ANG-1995) Andrew D. Farmer, Los Alamos National Laboratory, HIV Sequence Database, T-10, Mail Stop K710, Los Alamos, NM 87501, USA
location/Qualifiers
1..456
/organism="Human papillomavirus"
/mol_type="genomic DNA"
/isolate="OR 7632"
/db_xref="taxon:10566"
/note="collected by cervical lavage at Kaiser Permanente, Portland, Oregon, United States."
1..456
/gene="E6"
1..456
/gene="E6"
/codon_start=1
/product="early transforming protein E6"
/protein_id="AA91678.1"

gene
CDS
1..456
/gene="E6"
1..456
/gene="E6"
/codon_start=1
/product="early transforming protein E6"
/protein_id="AA91678.1"

FEATURES
source
Laboratory, HIV Sequence Database, T-10, Mail Stop K710, Los Alamos, NM 87501, USA
location/Qualifiers
1..456
/organism="Human papillomavirus"
/mol_type="genomic DNA"
/isolate="OR 8160"
/db_xref="taxon:10566"
/note="collected by cervical lavage at Kaiser Permanente, Portland, Oregon, United States."
1..456
/gene="E6"
1..456
/gene="E6"
/codon_start=1
/product="early transforming protein E6"
/protein_id="AA91679.1"

ORIGIN
Query Match 100.0%; Score 28; DB 10; Length 456;
Best Local Similarity 100.0%; Pred. No. 0.0025;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAACAGCAATACAAACCGTTGTG 28
|||||
265 GAACAGCAATACAAACCGTTGTG 292

Db 265 GAACAGCAATACAAACCGTTGTG 292

RESULT 49
HPU34132 456 bp DNA linear VRL 08-MAR-1996
LOCUS Human papillomavirus type 16, isolate OR 8160, early transforming
DEFINITION protein E6 (E6) gene, complete cds.
VERSION U34132.1 GI:1098769
KEYWORDS Human papillomavirus
SOURCE Human papillomavirus
ORGANISM Human papillomavirus
unclassified Papillomaviridae;
REFERENCE 1 (bases 1 to 456)
AUTHORS Yamada,T., Wheeler,C.M., Halpern,A.L., Stewart,A.C., Hildesheim,A. and Jensen,S.A.
TITLE Human papillomavirus type 16 variant lineages in United States populations characterized by nucleotide sequence analysis of the E6, E2, and L1 coding segments
JOURNAL U. Virol. 69 (12), 7743-7753 (1995)
PUBMED 7494284
REFERENCE 2 (bases 1 to 456)
AUTHORS Farmer,A.D.
TITLE Direct Submission
JOURNAL Submitted (16-ANG-1995) Andrew D. Farmer, Los Alamos National Laboratory, HIV Sequence Database, T-10, Mail Stop K710, Los Alamos, NM 87501, USA
location/Qualifiers
1..456
/organism="Human papillomavirus"
/mol_type="genomic DNA"
/isolate="OR 8160"
/db_xref="taxon:10566"
/note="collected by cervical lavage at Kaiser Permanente, Portland, Oregon, United States."
1..456
/gene="E6"
1..456
/gene="E6"
/codon_start=1
/product="early transforming protein E6"
/protein_id="AA91679.1"

ORIGIN
Query Match 100.0%; Score 28; DB 10; Length 456;
Best Local Similarity 100.0%; Pred. No. 0.0025;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAACAGCAATACAAACCGTTGTG 28
|||||
265 GAACAGCAATACAAACCGTTGTG 292

Db 265 GAACAGCAATACAAACCGTTGTG 292

RESULT 50
HPU34133

LOCUS HP034133 456 bp DNA linear VRL 08-MAR-1996
 DEFINITION Human papillomavirus type 16, isolate OR 8329, early transforming
 protein E6 (E6) gene, complete cds.
 ACCESSION U34133
 VERSION U34133.1 GI:1098771
 KEYWORDS
 SOURCE Human papillomavirus
 ORGANISM Human papillomavirus
 Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
 unclassified Papillomaviridae.
 REFERENCE 1 (bases 1 to 456)
 AUTHORS Yamada,T., Wheeler,C.M., Halpern,A.L., Stewart,A.C., Hildesheim,A.
 and Jenison,S.A.
 TITLE Human papillomavirus type 16 variant lineages in United States
 populations characterized by nucleotide sequence analysis of the
 E6, L2, and L1 coding segments
 JOURNAL J. VIROL. 69 (12), 7743-7753 (1995)
 PUBMED 7494284
 REFERENCE 2 (bases 1 to 456)
 AUTHORS Farmer,A.D.
 TITLE Direct Submission
 JOURNAL Submitted (16-NOV-1995) Andrew D. Farmer, Los Alamos National
 Laboratory, HIV Sequence Database, T-10, Mail Stop K710, Los
 Alamos, NM 87501, USA
 FEATURES
 source
 1. .456
 /organism="Human papillomavirus"
 /mol_type="genomic DNA"
 /isolate="OR 8329"
 /db_xref="taxon:10566"
 /note="collected by cervical lavage at Kaiser Permanente,
 Portland, Oregon, United States."
 1. .456
 /gene="E6"
 1. .456
 /gene="E6"
 /codon_start=1
 /product="early transforming protein E6"
 /protein_id="AA91580.1"
 /db_xref="GI:1098772"
 /translation="MPQDPQERPRKLPLQCTELQTTIHDIILLCVYCKQQLRREYVD
 FAFRDLCIVRDGNPYAVCDKCLTFYSKISEYHVCYSVYGTTLGQYNKPLCDLLIR
 CINQKPLCPBEKQRLDKKQRFHNIKRWYGRQWMSCCRSSRTREITOL"
 ORIGIN
 Query Match 100.0%; Score 28; DB 10; Length 456;
 Best Local Similarity 100.0%; Pred. No. 0.0025;
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GAACAGCAATACAAACCGTTGTGTG 28
 DB 265 GAACAGCAATACAAACCGTTGTGTG 292

Search completed: May 24, 2006, 06:53:50
 Job time : 915.046 secs

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 24, 2006, 05:46:37 ; Search time 638.284 Seconds
(without alignments)
2103.915 Million cell updates/sec

Title: US-10-601-913-85

Perfect score: 21

Sequence: 1 CAGGACACAGTGGCTTTGAC 21

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 6366136 seqs, 31973710525 residues

Total number of hits satisfying chosen parameters: 12732272

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 1000 summaries

Database :

GenBml:.*
1: gb_env:.*
2: gb_pat:.*
3: gb_ph:.*
4: gb_pl:.*
5: gb_pr:.*
6: gb_ro:.*
7: gb_ats:.*
8: gb_ey:.*
9: gb_un:.*
10: gb_vl:.*
11: gb_ov:.*
12: gb_htg:.*
13: gb_in:.*
14: gb_om:.*
15: gb_da:.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	21	100.0	21	2	AR349145 Sequence
2	21	100.0	21	2	AR349146 Sequence
3	21	100.0	21	2	AR349147 Sequence
4	21	100.0	21	2	AR349148 Sequence
5	21	100.0	104	2	DD187316 SPICEROSO
6	21	100.0	138	2	DD187319 SPICEROSO
7	21	100.0	170	10	HPUS9960 Human papil
8	21	100.0	245	10	HPUS9960 Human papil
9	21	100.0	253	10	HPUS9960 Human papil
10	21	100.0	271	10	HPUS9960 Human papil
11	21	100.0	272	10	HPUS9960 Human papil
12	21	100.0	273	10	HPUS9960 Human papil
13	21	100.0	274	10	HPUS9960 Human papil
14	21	100.0	298	10	HPUS9960 Human papil
15	21	100.0	370	10	HPUS9960 Human papil
16	21	100.0	421	10	HPUS9960 Human papil
17	21	100.0	451	10	HPUS9960 Human papil
18	21	100.0	451	10	HPUS9960 Human papil

19	21	100.0	451	10	AF404699 Human pap
20	21	100.0	451	10	AF404700 Human pap
21	21	100.0	451	10	AF404701 Human pap
22	21	100.0	451	10	AF404703 Human pap
23	21	100.0	456	2	AR167393 Sequence
24	21	100.0	456	2	AR167393 Sequence
25	21	100.0	456	10	AF327851 Human pap
26	21	100.0	456	10	AF327851 Human pap
27	21	100.0	456	10	AF327851 Human pap
28	21	100.0	456	10	AF327851 Human pap
29	21	100.0	456	10	AF327851 Human pap
30	21	100.0	456	10	AF327851 Human pap
31	21	100.0	456	10	AF327851 Human pap
32	21	100.0	456	10	AF327851 Human pap
33	21	100.0	456	10	AF327851 Human pap
34	21	100.0	456	10	AF327851 Human pap
35	21	100.0	456	10	AF327851 Human pap
36	21	100.0	456	10	AF327851 Human pap
37	21	100.0	456	10	AF327851 Human pap
38	21	100.0	456	10	AF327851 Human pap
39	21	100.0	456	10	AF327851 Human pap
40	21	100.0	456	10	AF327851 Human pap
41	21	100.0	456	10	AF327851 Human pap
42	21	100.0	456	10	AF327851 Human pap
43	21	100.0	456	10	AF327851 Human pap
44	21	100.0	456	10	AF327851 Human pap
45	21	100.0	456	10	AF327851 Human pap
46	21	100.0	456	10	AF327851 Human pap
47	21	100.0	456	10	AF327851 Human pap
48	21	100.0	456	10	AF327851 Human pap
49	21	100.0	456	10	AF327851 Human pap
50	21	100.0	456	10	AF327851 Human pap
51	21	100.0	456	10	AF327851 Human pap
52	21	100.0	456	10	AF327851 Human pap
53	21	100.0	456	10	AF327851 Human pap
54	21	100.0	456	10	AF327851 Human pap
55	21	100.0	456	10	AF327851 Human pap
56	21	100.0	456	10	AF327851 Human pap
57	21	100.0	456	10	AF327851 Human pap
58	21	100.0	456	10	AF327851 Human pap
59	21	100.0	456	10	AF327851 Human pap
60	21	100.0	456	10	AF327851 Human pap
61	21	100.0	456	10	AF327851 Human pap
62	21	100.0	456	10	AF327851 Human pap
63	21	100.0	456	10	AF327851 Human pap
64	21	100.0	456	10	AF327851 Human pap
65	21	100.0	456	10	AF327851 Human pap
66	21	100.0	456	10	AF327851 Human pap
67	21	100.0	456	10	AF327851 Human pap
68	21	100.0	456	10	AF327851 Human pap
69	21	100.0	456	10	AF327851 Human pap
70	21	100.0	456	10	AF327851 Human pap
71	21	100.0	456	10	AF327851 Human pap
72	21	100.0	456	10	AF327851 Human pap
73	21	100.0	456	10	AF327851 Human pap
74	21	100.0	456	10	AF327851 Human pap
75	21	100.0	456	10	AF327851 Human pap
76	21	100.0	456	10	AF327851 Human pap
77	21	100.0	456	10	AF327851 Human pap
78	21	100.0	456	10	AF327851 Human pap
79	21	100.0	456	10	AF327851 Human pap
80	21	100.0	456	10	AF327851 Human pap
81	21	100.0	456	10	AF327851 Human pap
82	21	100.0	456	10	AF327851 Human pap
83	21	100.0	456	10	AF327851 Human pap
84	21	100.0	456	10	AF327851 Human pap
85	21	100.0	456	10	AF327851 Human pap
86	21	100.0	456	10	AF327851 Human pap
87	21	100.0	456	10	AF327851 Human pap
88	21	100.0	456	10	AF327851 Human pap
89	21	100.0	456	10	AF327851 Human pap
90	21	100.0	456	10	AF327851 Human pap
91	21	100.0	456	10	AF327851 Human pap

C 92	21	100.0	477	10	AF486322	Human pap	C 165	21	100.0	1267	10	HPA388063	AJ388063	Human pap
C 93	21	100.0	477	10	AF486323	Human pap	C 166	21	100.0	1267	10	HPA388064	AJ388064	Human pap
C 94	21	100.0	477	10	AF486324	Human pap	C 167	21	100.0	1267	10	HPA388065	AJ388065	Human pap
C 95	21	100.0	477	10	AF486325	Human pap	C 168	21	100.0	1267	10	HPA388066	AJ388066	Human pap
C 96	21	100.0	477	10	DQ323982	Human pap	C 169	21	100.0	1267	10	HPA388067	AJ388067	Human pap
C 97	21	100.0	482	10	AF404658	Human pap	C 170	21	100.0	1267	10	HPA388068	AJ388068	Human pap
C 98	21	100.0	488	10	AF404702	Human pap	C 171	21	100.0	3531	10	AF001599	AF001599	Human pap
C 99	21	100.0	519	2	AR095319	Sequence	C 172	21	100.0	7902	2	AX800450	AX800450	Sequence
C 100	21	100.0	519	2	AR173472	Sequence	C 173	21	100.0	7904	2	BD070940	BD070940	Sequence
C 101	21	100.0	540	10	AF404705	Human pap	C 174	21	100.0	7904	2	BD225188	BD225188	Human pap
C 102	21	100.0	543	10	AF404653	Human pap	C 175	21	100.0	7904	2	CS073238	CS073238	Sequence
C 103	21	100.0	543	10	AF404657	Human pap	C 176	21	100.0	7904	2	CS195904	CS195904	Sequence
C 104	21	100.0	543	10	AF404706	Human pap	C 177	21	100.0	7904	2	I70142	I70142	Sequence
C 105	21	100.0	570	10	AY112663	Human pap	C 178	21	100.0	7904	2	I70143	I70143	Sequence
C 106	21	100.0	592	10	AY098923	Human pap	C 179	21	100.0	7904	2	I72345	I72345	Sequence
C 107	21	100.0	596	10	AY098919	Human pap	C 180	21	100.0	7904	2	AX763489	AX763489	Sequence
C 108	21	100.0	598	2	CO410317	Sequence	C 181	21	100.0	7904	10	AF125673	AF125673	Human pap
C 109	21	100.0	600	10	AY098922	Human pap	C 182	21	100.0	7904	10	AF472509	AF472509	Human pap
C 110	21	100.0	601	10	AY098918	Human pap	C 183	21	100.0	7904	10	AF536179	AF536179	Human pap
C 111	21	100.0	602	10	AY112662	Human pap	C 184	21	100.0	7904	10	PPH16	PPH16	Human pap
C 112	21	100.0	654	10	AF548023	Human pap	C 185	21	100.0	7905	10	AF534061	AF534061	Human pap
C 113	21	100.0	693	10	AF469198	Human pap	C 186	21	100.0	7905	10	HPH89348	HPH89348	Human pap
C 114	21	100.0	741	10	AF469197	Human pap	C 187	21	100.0	7906	10	AF536180	AF536180	Human pap
C 115	21	100.0	743	2	CO414264	Sequence	C 188	21	100.0	7906	10	AF686579	AF686579	Human pap
C 116	21	100.0	745	2	AX188300	Sequence	C 189	21	100.0	7906	10	AF686580	AF686580	Human pap
C 117	21	100.0	755	10	AY089951	Human pap	C 190	21	100.0	7906	10	AY686581	AY686581	Human pap
C 118	21	100.0	755	10	AY089952	Human pap	C 191	21	100.0	7906	10	AY686584	AY686584	Human pap
C 119	21	100.0	755	10	AY089953	Human pap	C 192	21	100.0	7907	10	AY686582	AY686582	Human pap
C 120	21	100.0	755	10	AY089954	Human pap	C 193	21	100.0	7907	10	AY686583	AY686583	Human pap
C 121	21	100.0	755	10	AY089955	Human pap	C 194	21	100.0	7908	10	AF472508	AF472508	Human pap
C 122	21	100.0	776	2	I19491	Sequence	C 195	21	95.2	20	2	A21235	A21235	Sequence
C 123	21	100.0	779	2	AR199235	Sequence	C 196	21	95.2	20	2	BD236838	BD236838	Sequence
C 124	21	100.0	780	10	AF187867	Human pap	C 197	21	95.2	20	2	CO814785	CO814785	Sequence
C 125	21	100.0	780	10	AF187867	Human pap	C 198	21	95.2	20	2	AR267070	AR267070	Sequence
C 126	21	100.0	780	10	AF187868	Human pap	C 199	21	95.2	20	2	AR652060	AR652060	Sequence
C 127	21	100.0	780	10	AF187869	Human pap	C 200	21	95.2	20	2	AX030354	AX030354	Sequence
C 128	21	100.0	790	2	I88813	Sequence	C 201	21	95.2	20	2	AX379361	AX379361	Sequence
C 129	21	100.0	801	2	AR095318	Sequence	C 202	21	95.2	21	2	DD178908	DD178908	METHOD
C 130	21	100.0	801	2	AR173471	Sequence	C 203	21	95.2	21	2	AX802808	AX802808	Sequence
C 131	21	100.0	822	2	A98742	Sequence	C 204	19.4	92.4	477	10	AF486319	AF486319	Human pap
C 132	21	100.0	822	2	BD080312	Vaccine.	C 205	19.4	92.4	7906	10	AF402578	AF402578	Human pap
C 133	21	100.0	822	2	BD103161	Vaccine.	C 206	19	90.5	20	2	E05240	E05240	Part of DNA
C 134	21	100.0	822	2	AR183573	Sequence	C 207	19	90.5	20	2	AX805231	AX805231	Sequence
C 135	21	100.0	822	2	AX020926	Sequence	C 208	19	90.5	29	2	AR142033	AR142033	Sequence
C 136	21	100.0	879	2	A98748	Sequence	C 209	19	90.5	29	2	I59900	I59900	Sequence
C 137	21	100.0	879	2	BD080315	Vaccine.	C 210	19	90.5	29	2	I86758	I86758	Sequence
C 138	21	100.0	879	2	BD103164	Vaccine.	C 211	19	90.5	29	2	I95783	I95783	Sequence
C 139	21	100.0	879	2	AR183576	Sequence	C 212	19	90.5	29	2	AR708762	AR708762	Sequence
C 140	21	100.0	879	2	AX020933	Sequence	C 213	19	90.5	194799	5	AC006430	AC006430	Homo sapi
C 141	21	100.0	921	10	HPA388069	Human pap	C 214	18.4	87.6	89	2	DD187320	DD187320	SPLICING
C 142	21	100.0	939	2	AX403959	Sequence	C 215	18.4	87.6	799	5	AY650345	AY650345	Macaca fa
C 143	21	100.0	939	2	AX460903	Sequence	C 216	18.4	87.6	1596	2	CO731442	CO731442	Sequence
C 144	21	100.0	1000	2	A07622	Amplified	C 217	18.4	87.6	2317	2	CO719194	CO719194	Sequence
C 145	21	100.0	1000	2	I18920	Sequence	C 218	18.4	87.6	2345	2	CO824165	CO824165	Sequence
C 146	21	100.0	1116	2	A98744	Sequence	C 219	18.4	87.6	2545	5	HDWCABA	HDWCABA	Sequence
C 147	21	100.0	1116	2	BD080313	Vaccine.	C 220	18.4	87.6	2550	2	I34431	I34431	Homo sapien
C 148	21	100.0	1116	2	BD103162	Vaccine.	C 221	18.4	87.6	2550	2	I46829	I46829	Sequence
C 149	21	100.0	1116	2	AR183574	Sequence	C 222	18.4	87.6	2550	2	I55098	I55098	Sequence
C 150	21	100.0	1116	2	AX020928	Sequence	C 223	18.4	87.6	2742	2	CS162382	CS162382	Sequence
C 151	21	100.0	1154	10	PPH16E	Human pap	C 224	18.4	87.6	2742	5	AB035305	AB035305	Homo sapi
C 152	21	100.0	1173	2	A98752	Sequence	C 225	18.4	87.6	2812	6	MMCCAD8	MMCCAD8	Homo sapi
C 153	21	100.0	1173	2	BD080317	Vaccine.	C 226	18.4	87.6	2929	6	AX933856	AX933856	Sequence
C 154	21	100.0	1173	2	BD103166	Vaccine.	C 227	18.4	87.6	3126	6	AB010436	AB010436	Rattus ra
C 155	21	100.0	1173	2	AR183578	Sequence	C 228	18.4	87.6	3136	2	I34425	I34425	Sequence
C 156	21	100.0	1173	2	AX020936	Sequence	C 229	18.4	87.6	3136	2	I55092	I55092	Sequence
C 157	21	100.0	1267	10	HPA242681	Human pap	C 230	18.4	87.6	3502	5	AK124734	AK124734	Homo sapi
C 158	21	100.0	1267	10	HPA388056	Human pap	C 231	18.4	87.6	3552	6	BC057581	BC057581	Mus muscu
C 159	21	100.0	1267	10	HPA388057	Human pap	C 232	18.4	87.6	4334	6	BC060200	BC060200	Mus muscu
C 160	21	100.0	1267	10	HPA388058	Human pap	C 233	18.4	87.6	138810	6	AC103397	AC103397	Mus muscu
C 161	21	100.0	1267	10	HPA388059	Human pap	C 234	18.4	87.6	171239	12	AC010652	AC010652	Homo sapi
C 162	21	100.0	1267	10	HPA388060	Human pap	C 235	18.4	87.6	173905	5	AC092125	AC092125	Homo sapi
C 163	21	100.0	1267	10	HPA388061	Human pap	C 236	18.4	87.6	177751	12	AC016844	AC016844	Homo sapi
C 164	21	100.0	1267	10	HPA388062	Human pap	C 237	18.4	87.6	190333	5	AC092129	AC092129	Homo sapi

C 238	18.4	87.6	193488	5	AL136097	AL136097 Human DNA	311	17	81.0	168706	12	AC016151	AC016151 Homo sapi
C 239	18.4	87.6	198116	5	AC130217	Mus muscu	C 312	17	81.0	175223	5	AC009806	AC009806 Homo sapi
C 240	18.4	87.6	203790	5	AC010422	Homo sapi	313	17	81.0	178620	6	AC103359	AC103359 Mus muscu
C 241	18.4	87.6	212437	12	AL354924	Homo sapi	314	17	81.0	181467	6	AL844586	AL844586 Mouse DNA
C 242	18.4	87.6	213528	12	AC133688	Rattus no	315	17	81.0	192695	6	CT025537	CT025537 Mouse DNA
C 243	18.4	87.6	213770	12	AC103152	Rattus no	C 316	17	81.0	194056	5	AL5590708	AL5590708 Human DNA
C 244	18.4	87.6	224135	12	AC117143	Rattus no	C 317	17	81.0	198549	6	AC162903	AC162903 Mus muscu
C 245	18.4	87.6	251205	12	AC112454	Rattus no	318	17	81.0	204251	6	AC108430	AC108430 Mus muscu
C 246	18.4	87.6	272335	12	AC127222	Rattus no	319	17	81.0	206439	12	AC160069	AC160069 Bos tauru
C 247	18.4	87.6	340625	12	AC117847	Rattus no	C 320	17	81.0	209901	12	AC102707	AC102707 Mus muscu
C 248	18.4	87.7	99	2	DD187318	SPLICROSO	321	17	81.0	215609	12	AC170573	AC170573 Bos tauru
C 249	17.8	84.8	610	7	BV341640	SV341640 S230P6476	C 322	17	81.0	239312	12	AC111457	AC111457 Rattus no
C 250	17.8	84.8	766	2	AX186489	Sequence	C 323	17	81.0	247084	12	AC125034	AC125034 Mus muscu
C 251	17.8	84.8	1546	6	BC109362	BC109362 Mus muscu	C 324	17	81.0	254973	12	AC127437	AC127437 Rattus no
C 252	17.8	84.8	2002	6	BC109363	BC109363 Mus muscu	C 325	17	81.0	253936	12	AC094412	AC094412 Rattus no
C 253	17.8	84.8	99601	12	AL160402	AL160402 Homo sapi	326	17	81.0	270241	12	AC112064	AC112064 Rattus no
C 254	17.8	84.8	99686	12	AC002518	AC002518 Homo sapi	C 327	17	81.0	298446	12	AC117043	AC117043 Rattus no
C 255	17.8	84.8	120101	5	AC064872	AC064872 Homo sapi	C 328	16.8	80.0	105	2	CQ423136	CQ423136 Sequence
C 256	17.8	84.8	127917	2	CO861449	CO861449 Sequence	329	16.8	80.0	201	7	BV201394	BV201394 sqm20648
C 257	17.8	84.8	127917	5	HS257A7	HS257A7 Human DNA	330	16.8	80.0	312	2	CO731534	CO731534 Sequence
C 258	17.8	84.8	127917	5	AC009239	AC009239 Homo sapi	331	16.8	80.0	325	2	AX987374	AX987374 Sequence
C 259	17.8	84.8	147650	12	CT025675	CT025675 Mus muscu	332	16.8	80.0	325	2	BD122233	BD122233 EST and e
C 260	17.8	84.8	153016	12	AC173953	AC173953 Sorex ara	333	16.8	80.0	325	2	AR426680	AR426680 Sequence
C 261	17.8	84.8	154573	12	AC148733	AC148733 Callithrix	334	16.8	80.0	331	2	CO731143	CO731143 Sequence
C 262	17.8	84.8	155305	12	AC117902	AC117902 Rattus no	335	16.8	80.0	408	2	AX552180	AX552180 Sequence
C 263	17.8	84.8	151875	5	AC171740	AC171740 Mustela f	336	16.8	80.0	415	2	BD034459	BD034459 Sequence
C 264	17.8	84.8	163976	5	AC108483	AC108483 Homo sapi	337	16.8	80.0	415	2	AR734720	AR734720 Sequence
C 265	17.8	84.8	163998	5	AC022166	AC022166 Homo sapi	338	16.8	80.0	415	2	AX898926	AX898926 Sequence
C 266	17.8	84.8	164347	12	AC119669	AC119669 Rattus no	339	16.8	80.0	450	2	AX972653	AX972653 Sequence
C 267	17.8	84.8	164519	5	AL138712	AL138712 Human DNA	340	16.8	80.0	450	2	BD111372	BD111372 EST and e
C 268	17.8	84.8	166955	12	AC080161	AC080161 Homo sapi	341	16.8	80.0	450	2	AR415819	AR415819 Sequence
C 269	17.8	84.8	197832	12	AC019332	AC019332 Homo sapi	342	16.8	80.0	492	2	AX985368	AX985368 Sequence
C 270	17.8	84.8	201067	6	AC102392	AC102392 Mus muscu	343	16.8	80.0	492	2	BD120227	BD120227 EST and e
C 271	17.8	84.8	201094	12	AC020995	AC020995 Homo sapi	344	16.8	80.0	492	2	AR424674	AR424674 Sequence
C 272	17.8	84.8	212178	12	AC125925	AC125925 Rattus no	345	16.8	80.0	495	2	BD109009	BD109009 EST and e
C 273	17.8	84.8	212854	6	AC140674	AC140674 Mus muscu	346	16.8	80.0	495	2	AX970290	AX970290 Sequence
C 274	17.8	84.8	224450	12	AC016311	AC016311 Homo sapi	347	16.8	80.0	495	2	AR413456	AR413456 Sequence
C 275	17.8	84.8	242745	12	AC109521	AC109521 Rattus no	348	16.8	80.0	497	2	BD108056	BD108056 Sequence
C 276	17.8	84.8	247586	12	AC102960	AC102960 Rattus no	349	16.8	80.0	497	2	AX108035	AX108035 EST and e
C 277	17.8	84.8	250284	12	AC106077	AC106077 Rattus no	350	16.8	80.0	497	2	AR412503	AR412503 Sequence
C 278	17.8	84.8	256733	12	AC111836	AC111836 Rattus no	351	16.8	80.0	499	2	AX428904	AX428904 Sequence
C 279	17.4	82.9	371	2	CO529827	CO529827 Sequence	352	16.8	80.0	499	2	AX430235	AX430235 Sequence
C 280	17.4	82.9	13015	15	AP435074	AP435074 Corynebac	353	16.8	80.0	501	2	AX969346	AX969346 Sequence
C 281	17.4	82.9	72788	4	NCB2A19	NCB2A19 Neurospor	354	16.8	80.0	501	2	BD108065	BD108065 EST and e
C 282	17.4	82.9	101954	11	BX942844	BX942844 Zebrafish	355	16.8	80.0	501	2	AR412512	AR412512 Sequence
C 283	17.4	82.9	146757	11	BX248512	BX248512 Zebrafish	356	16.8	80.0	513	2	AX970291	AX970291 Sequence
C 284	17.4	82.9	154034	11	BX927371	BX927371 Zebrafish	357	16.8	80.0	513	2	BD109010	BD109010 EST and e
C 285	17.4	82.9	162758	12	CT025756	CT025756 Danio rer	358	16.8	80.0	513	2	AR413457	AR413457 Sequence
C 286	17.4	82.9	165669	14	AC127465	AC127465 Didelphis	359	16.8	80.0	518	2	CO262874	CO262874 Sequence
C 287	17.4	82.9	169604	6	AC107775	AC107775 Mus muscu	360	16.8	80.0	518	2	CQ299917	CQ299917 Sequence
C 288	17.4	82.9	170731	6	AC115846	AC115846 Mus muscu	361	16.8	80.0	524	7	BV518499	BV518499 srx37BD04.
C 289	17.4	82.9	181261	6	AC133346	AC133346 Mus muscu	C 362	16.8	80.0	566	7	HSC19C12	AL158580 H. sapiens
C 290	17.4	82.9	194704	6	AC107772	AC107772 Mus muscu	363	16.8	80.0	568	7	BV001968	BV001968 S208P6073
C 291	17.4	82.9	196501	5	AC005908	AC005908 Homo sapi	364	16.8	80.0	570	2	AX552179	AX552179 Sequence
C 292	17.4	82.9	202726	11	AL773593	AL773593 Zebrafish	365	16.8	80.0	571	2	AX972654	AX972654 Sequence
C 293	17.4	82.9	202751	12	BX248237	BX248237 Danio rer	366	16.8	80.0	571	2	BD111373	BD111373 EST and e
C 294	17.4	82.9	217927	6	AC159308	AC159308 Mus muscu	367	16.8	80.0	571	2	AR415820	AR415820 Sequence
C 295	17.4	82.9	237007	6	AL627166	AL627166 Mouse DNA	368	16.8	80.0	572	2	AX971891	AX971891 Sequence
C 296	17.4	82.9	244908	12	AC130992	AC130992 Rattus no	369	16.8	80.0	572	2	BD110610	BD110610 EST and e
C 297	17.4	82.9	246968	12	AC131630	AC131630 Rattus no	370	16.8	80.0	572	2	AR415057	AR415057 Sequence
C 298	17.4	82.9	256103	12	AC129636	AC129636 Rattus no	371	16.8	80.0	579	2	CO225389	CO225389 Sequence
C 299	17.4	82.9	277252	12	AC175278	AC175278 Bos tauru	372	16.8	80.0	610	5	BC056684	BC056684 Homo sapi
C 300	17.4	82.9	291200	12	AC151529	AC151529 Bos tauru	C 373	16.8	80.0	636	5	BV180243	BV180243 equm10855
C 301	17.4	82.9	292398	12	AC121748	AC121748 Rattus no	374	16.8	80.0	643	5	BC103985	BC103985 Homo sapi
C 302	17	81.0	834	11	DQ096969	DQ096969 Xenopus 1	375	16.8	80.0	643	5	BC103986	BC103986 Homo sapi
C 303	17	81.0	2421	11	BC073467	BC073467 Xenopus 1	376	16.8	80.0	643	5	BC103987	BC103987 Homo sapi
C 304	17	81.0	110000	5	AL3094437	AL3094437 Rattus no	377	16.8	80.0	655	5	HUMZB16C11	AL086653 Homo sapi
C 305	17	81.0	146853	12	AL136527	AL136527 Human DNA	378	16.8	80.0	671	5	AK130974	AK130974 Homo sapi
C 306	17	81.0	147252	12	AC182174	AC182174 Bos tauru	379	16.8	80.0	676	5	AX060629	AX060629 Sequence
C 307	17	81.0	152358	12	AC018379	AC018379 Homo sapi	380	16.8	80.0	704	2	BC062786	BC062786 Homo sapi
C 308	17	81.0	162381	5	AC074045	AC074045 Homo sapi	381	16.8	80.0	720	2	BD147498	BD147498 Primer fo
C 309	17	81.0	163310	12	AC016147	AC016147 Homo sapi	382	16.8	80.0	720	2	AX867436	AX867436 Sequence
C 310	17	81.0	165415	5	AL359181	AL359181 Human DNA	383	16.8	80.0	755	2	BD147493	BD147493 Primer fo

384	16.8	80.0	755	2	AX867431	Sequence	C 457	16.8	80.0	131111	6	AC140781	Mus muscu
385	16.8	80.0	760	7	BV549379	BD150259	C 458	16.8	80.0	131398	5	HS445C9	295115 Human DNA s
C 386	16.8	80.0	767	2	BD150259	AX870197	C 459	16.8	80.0	133852	5	AC010326	AC096845 Takifugu
C 387	16.8	80.0	767	2	AX870197	AX870197	C 460	16.8	80.0	134105	11	AC096845	AX870197
388	16.8	80.0	780	13	AX870197	AX870197	C 461	16.8	80.0	136247	12	BSX42845	BSX42845
389	16.8	80.0	788	2	AX876239	AX876239	C 462	16.8	80.0	142003	5	AC098800	AC098800
C 390	16.8	80.0	808	7	BV517049	BV517049	C 463	16.8	80.0	144087	5	AP001621	AP001621 Homo sapi
C 391	16.8	80.0	1000	7	BV554334	BV554334	C 464	16.8	80.0	145108	5	AL683886	AL683886 Human DNA
C 392	16.8	80.0	1075	2	AX540436	AX540436	C 465	16.8	80.0	146699	11	BS000226	BS000226 Pan trogl
393	16.8	80.0	1144	2	AR542136	AR542136	C 466	16.8	80.0	147881	11	BSX42831	BSX42831 Zebrafish
394	16.8	80.0	1144	2	AR692928	AR692928	C 467	16.8	80.0	147902	11	BSX48026	BSX48026 Zebrafish
C 395	16.8	80.0	1339	2	CQ723387	CQ723387	C 468	16.8	80.0	148385	12	AC166243	AC166243 Mus muscu
C 396	16.8	80.0	1637	5	BC067277	BC067277	C 469	16.8	80.0	150763	6	AC140785	AC140785 Mus muscu
397	16.8	80.0	1659	5	BC015312	BC015312	C 470	16.8	80.0	150872	12	AC148037	AC148037 Rattus no
C 398	16.8	80.0	1905	2	BD157905	BD157905	C 471	16.8	80.0	152563	12	AC175211	AC175211 Echinos
399	16.8	80.0	1905	2	AX879566	AX879566	C 472	16.8	80.0	154632	6	AC131702	AC131702 Mus muscu
400	16.8	80.0	1905	2	AK022668	AK022668	C 473	16.8	80.0	156194	12	AC022806	AC022806 Homo sapi
401	16.8	80.0	2015	2	AX399227	AX399227	C 474	16.8	80.0	157485	6	AC121991	AC121991 Mus muscu
402	16.8	80.0	2070	2	CS168404	CS168404	C 475	16.8	80.0	157768	12	AC175438	AC175438 Bos tauru
403	16.8	80.0	2070	2	CS239182	CS239182	C 476	16.8	80.0	158638	6	AC133087	AC133087 Mus muscu
404	16.8	80.0	2070	2	AX834525	AX834525	C 477	16.8	80.0	159027	12	AC152029	AC152029 Daaypus n
405	16.8	80.0	2070	5	AK097041	AK097041	C 478	16.8	80.0	159451	5	AC012618	AC012618 Homo sapi
406	16.8	80.0	2092	2	BD157900	BD157900	C 479	16.8	80.0	159451	5	AC012618	AC012618 Homo sapi
407	16.8	80.0	2092	2	AX879558	AX879558	C 480	16.8	80.0	159962	6	AC154229	AC154229 Bos tauru
408	16.8	80.0	2092	5	BC022663	BC022663	C 481	16.8	80.0	160169	12	AC051664	AC051664 Homo sapi
409	16.8	80.0	2136	5	AF129264	AF129264	C 482	16.8	80.0	161198	12	AC158677	AC158677 Homo sapi
C 410	16.8	80.0	2139	5	AB179282	AB179282	C 483	16.8	80.0	161258	12	AC158727	AC158727 Rhinoloph
411	16.8	80.0	2190	5	AB179282	AB179282	C 484	16.8	80.0	163915	12	AC087451	AC087451 Homo sapi
412	16.8	80.0	2468	2	AR718892	AR718892	C 485	16.8	80.0	164789	6	AC147185	AC147185 Mus muscu
413	16.8	80.0	2468	2	AX747867	AX747867	C 486	16.8	80.0	165093	6	AC069444	AC069444 Homo sapi
414	16.8	80.0	2468	5	AK093030	AK093030	C 487	16.8	80.0	165671	12	ALB84163	ALB84163 Mouse DNA
415	16.8	80.0	2574	5	BC041820	BC041820	C 488	16.8	80.0	166116	6	ALB84163	ALB84163 Mouse DNA
416	16.8	80.0	2742	5	AB110214	AB110214	C 489	16.8	80.0	166244	5	AL365222	AL365222 Homo sapi
417	16.8	80.0	2779	5	HSM80634	HSM80634	C 490	16.8	80.0	166687	5	AP003733	AP003733 Homo sapi
C 418	16.8	80.0	2906	5	AF129263	AF129263	C 491	16.8	80.0	167922	12	AC014605	AC014605 Homo sapi
419	16.8	80.0	3113	2	CQ849816	CQ849816	C 492	16.8	80.0	168462	11	EX005308	EX005308 Zebrafish
420	16.8	80.0	3113	5	AK126871	AK126871	C 493	16.8	80.0	168540	12	AC129044	AC129044 Rattus no
C 421	16.8	80.0	3233	2	BD160187	BD160187	C 494	16.8	80.0	170121	5	AF064860	AF064860 Homo sapi
C 422	16.8	80.0	3233	2	AX883204	AX883204	C 495	16.8	80.0	171014	6	AC154215	AC154215 Mus muscu
C 423	16.8	80.0	3233	5	AK023769	AK023769	C 496	16.8	80.0	172924	12	AC167304	AC167304 Homo sapi
C 424	16.8	80.0	3579	5	AK123008	AK123008	C 497	16.8	80.0	173058	5	AC104564	AC104564 Homo sapi
425	16.8	80.0	3840	14	AF251797	AF251797	C 498	16.8	80.0	173152	6	AC142168	AC142168 Mus muscu
426	16.8	80.0	5317	5	HSM802401	HSM802401	C 499	16.8	80.0	173589	12	AC019276	AC019276 Mus muscu
427	16.8	80.0	8693	14	OCWAPG	OCWAPG	C 500	16.8	80.0	174305	12	AC074126	AC074126 Homo sapi
428	16.8	80.0	35705	5	AC010522	AC010522	C 501	16.8	80.0	175400	6	AC170178	AC170178 Mus muscu
429	16.8	80.0	35711	5	AC020951	AC020951	C 502	16.8	80.0	175759	6	AF220294	AF220294 Mus muscu
430	16.8	80.0	38653	5	AR285078	AR285078	C 503	16.8	80.0	176140	5	AC092043	AC092043 Homo sapi
431	16.8	80.0	39880	5	AC010645	AC010645	C 504	16.8	80.0	176532	12	AC161824	AC161824 Mus muscu
C 432	16.8	80.0	46651	5	AC000404	AC000404	C 505	16.8	80.0	176765	12	AC025376	AC025376 Homo sapi
C 433	16.8	80.0	53242	5	AL365494	AL365494	C 506	16.8	80.0	178184	5	AC008770	AC008770 Homo sapi
C 434	16.8	80.0	53945	5	AL139111	AL139111	C 507	16.8	80.0	178416	6	AC127554	AC127554 Mus muscu
C 435	16.8	80.0	64931	5	AC128716	AC128716	C 508	16.8	80.0	179272	12	AC069149	AC069149 Homo sapi
436	16.8	80.0	70474	11	BSX322786	BSX322786	C 509	16.8	80.0	179272	12	AC069149	AC069149 Homo sapi
437	16.8	80.0	70474	11	BSX322786	BSX322786	C 510	16.8	80.0	179561	12	AC009397	AC009397 Homo sapi
438	16.8	80.0	71937	12	AC100198	AC100198	C 511	16.8	80.0	180076	12	AC022177	AC022177 Homo sapi
439	16.8	80.0	72354	5	AL357123	AL357123	C 512	16.8	80.0	180410	6	AC161867	AC161867 Mus muscu
C 440	16.8	80.0	73434	12	AC123783	AC123783	C 513	16.8	80.0	180510	5	AC022415	AC022415 Homo sapi
C 441	16.8	80.0	75388	11	BSX36441	BSX36441	C 514	16.8	80.0	181714	12	AC151893	AC151893 Aotus tan
442	16.8	80.0	84613	5	AL136140	AL136140	C 515	16.8	80.0	183390	12	AC162536	AC162536 Bos tauru
443	16.8	80.0	84876	5	BSX284698	BSX284698	C 516	16.8	80.0	185079	12	AC135702	AC135702 Rattus no
C 444	16.8	80.0	86081	5	AF364606	AF364606	C 517	16.8	80.0	185800	12	AC079366	AC079366 Mus muscu
445	16.8	80.0	94851	5	AC004923	AC004923	C 518	16.8	80.0	186142	12	AL390124	AL390124 Homo sapi
446	16.8	80.0	105238	5	AC011458	AC011458	C 519	16.8	80.0	186595	6	AC140053	AC140053 Mus muscu
C 447	16.8	80.0	106648	12	AC004588	AC004588	C 520	16.8	80.0	186803	12	AC151868	AC151868 Callithr
448	16.8	80.0	110000	12	AL954350	AL954350	C 521	16.8	80.0	187110	6	AC131303	AC131303 Mus muscu
C 449	16.8	80.0	110000	12	AL954350	AL954350	C 522	16.8	80.0	187415	12	AC026352	AC026352 Homo sapi
450	16.8	80.0	110000	15	CP000075	CP000075	C 523	16.8	80.0	187734	11	AL928657	AL928657 Zebrafish
C 451	16.8	80.0	110000	15	CR522870	CR522870	C 524	16.8	80.0	187947	12	AC162369	AC162369 Mus muscu
C 452	16.8	80.0	110000	15	AB015928	AB015928	C 525	16.8	80.0	188269	12	AC170183	AC170183 Homo sapi
C 453	16.8	80.0	110000	15	BA000040	BA000040	C 526	16.8	80.0	188346	12	AC010189	AC010189 Homo sapi
C 454	16.8	80.0	110746	12	CT027669	CT027669	C 527	16.8	80.0	188570	12	AC110390	AC110390 Rattus no
455	16.8	80.0	114670	11	BSX36320	BSX36320	C 528	16.8	80.0	188734	12	AC069443	AC069443 Homo sapi
456	16.8	80.0	120376	11	CR381573	CR381573	C 529	16.8	80.0	190626	12	AC062036	AC062036 Homo sapi

C 530	16.8	80.0	190632	6	AC110622	AC110622 Mus muscu	C 603	16.8	80.0	238625	11	BX005057	BX005057 Zebrafish
C 531	16.8	80.0	190802	6	AC124424	AC124424 Mus muscu	C 604	16.8	80.0	240000	12	AC006393	AC006393 Homo sapi
C 532	16.8	80.0	191038	6	AC165327	AC165327 Mus muscu	C 605	16.8	80.0	240774	12	AC130391	AC130391 Rattus no
C 533	16.8	80.0	191222	12	AC022058	AC022058 Homo sapi	C 606	16.8	80.0	243689	12	AC129281	AC129281 Rattus no
C 534	16.8	80.0	191847	12	CR735111	CR735111 Danio rer	C 607	16.8	80.0	244201	12	AC098519	AC098519 Rattus no
C 535	16.8	80.0	192219	5	RP43002119	RP43002119 Pan trogl	C 608	16.8	80.0	244961	12	AC098519	AC098519 Rattus no
C 536	16.8	80.0	194257	11	BX255948	BX255948 Zebrafish	C 609	16.8	80.0	244574	12	AC166377	AC166377 Bos tauru
C 537	16.8	80.0	195006	6	AC121571	AC121571 Mus muscu	C 610	16.8	80.0	245594	6	AC103190	AC103190 Rattus no
C 538	16.8	80.0	195609	6	AC159127	AC159127 Mus muscu	C 611	16.8	80.0	246289	12	AC165426	AC165426 Mus muscu
C 539	16.8	80.0	197194	12	AL954690	AL954690 Mus muscu	C 612	16.8	80.0	246289	12	AC094407	AC094407 Rattus no
C 540	16.8	80.0	198900	6	AC123556	AC123556 Mus muscu	C 613	16.8	80.0	248748	12	AC130614	AC130614 Rattus no
C 541	16.8	80.0	199032	12	AC171843	AC171843 Bos tauru	C 614	16.8	80.0	250735	12	AC161693	AC161693 Bos tauru
C 542	16.8	80.0	199077	12	CR753886	CR753886 Danio rer	C 615	16.8	80.0	250810	12	AC115157	AC115157 Rattus no
C 543	16.8	80.0	199583	6	AC110205	AC110205 Mus muscu	C 616	16.8	80.0	252138	12	AC151886	AC151886 Salimtri b
C 544	16.8	80.0	199848	6	AL512647	AL512647 Mus muscu	C 617	16.8	80.0	253172	12	AC137356	AC137356 Rattus no
C 545	16.8	80.0	201266	12	AC116240	AC116240 Rattus no	C 618	16.8	80.0	253744	6	AC105147	AC105147 Rattus no
C 546	16.8	80.0	201300	5	AL137073	AL137073 Human DNA	C 619	16.8	80.0	259178	12	AC108971	AC108971 Rattus no
C 547	16.8	80.0	202133	5	AC092436	AC092436 Homo sapi	C 620	16.8	80.0	267392	12	AC099071	AC099071 Rattus no
C 548	16.8	80.0	203331	6	AC131752	AC131752 Mus muscu	C 621	16.8	80.0	273535	12	CT573447	CT573447 Danio rer
C 549	16.8	80.0	203336	6	AL591433	AL591433 Mouse DNA	C 622	16.8	80.0	275618	12	AC103413	AC103413 Rattus no
C 550	16.8	80.0	204006	12	AC170435	AC170435 Bos tauru	C 623	16.8	80.0	275710	12	AC135276	AC135276 Rattus no
C 551	16.8	80.0	204091	6	AC138368	AC138368 Mus muscu	C 624	16.8	80.0	281117	12	AC110645	AC110645 Rattus no
C 552	16.8	80.0	204825	6	AC158514	AC158514 Mus muscu	C 625	16.8	80.0	293507	12	AC097292	AC097292 Rattus no
C 553	16.8	80.0	205318	6	AC122894	AC122894 Mus muscu	C 626	16.8	80.0	293758	12	AC173816	AC173816 Bos tauru
C 554	16.8	80.0	205638	5	AC090018	AC090018 Homo sapi	C 627	16.8	80.0	297007	12	AC166528	AC166528 Bos tauru
C 555	16.8	80.0	206192	5	AC034259	AC034259 Homo sapi	C 628	16.8	80.0	298705	12	AC156690	AC156690 Bos tauru
C 556	16.8	80.0	207148	12	CT027779	CT027779 Danio rer	C 629	16.8	80.0	304570	12	AC167052	AC167052 Bos tauru
C 557	16.8	80.0	207223	6	AL733521	AL733521 Mouse DNA	C 630	16.8	80.0	312697	12	AL683889	AL683889 Homo sapi
C 558	16.8	80.0	207698	5	AC114491	AC114491 Homo sapi	C 631	16.8	80.0	312773	12	AC117017	AC117017 Rattus no
C 559	16.8	80.0	209196	12	AC017024	AC017024 Homo sapi	C 632	16.8	80.0	320074	12	AC162218	AC162218 Bos tauru
C 560	16.8	80.0	209365	6	AC121148	AC121148 Mus muscu	C 633	16.8	80.0	334392	12	AC118975	AC118975 Rattus no
C 561	16.8	80.0	209384	12	AC151295	AC151295 Mus muscu	C 634	16.8	80.0	336578	5	AP001745	AP001745 Homo sapi
C 562	16.8	80.0	210115	6	AC154340	AC154340 Mus muscu	C 635	16.8	80.0	340000	5	HS21C080	HS21C080 Homo sapi
C 563	16.8	80.0	210784	6	AC141564	AC141564 Mus muscu	C 636	16.8	80.0	342684	12	AC094255	AC094255 Rattus no
C 564	16.8	80.0	210968	12	AC123209	AC123209 Rattus no	C 637	16.8	80.0	345616	12	AC120730	AC120730 Rattus no
C 565	16.8	80.0	211176	12	AC165743	AC165743 Bos tauru	C 638	16.4	78.1	247	2	AR206466	AR206466 Sequence
C 566	16.8	80.0	211182	12	AC165470	AC165470 Bos tauru	C 639	16.4	78.1	247	2	AX147224	AX147224 Sequence
C 567	16.8	80.0	211901	6	AL772310	AL772310 Mouse DNA	C 640	16.4	78.1	364	2	CQ476429	CQ476429 Sequence
C 568	16.8	80.0	212253	12	AC140184	AC140184 Mus muscu	C 641	16.4	78.1	387	2	CQ398510	CQ398510 Sequence
C 569	16.8	80.0	213047	12	AC020557	AC020557 Homo sapi	C 642	16.4	78.1	387	2	CQ404797	CQ404797 Sequence
C 570	16.8	80.0	213086	12	AC119818	AC119818 Mus muscu	C 643	16.4	78.1	431	2	BD024558	BD024558 Sequence
C 571	16.8	80.0	213817	12	AC123649	AC123649 Mus muscu	C 644	16.4	78.1	431	2	AR724819	AR724819 Sequence
C 572	16.8	80.0	214765	6	AL731790	AL731790 Mouse DNA	C 645	16.4	78.1	431	2	AX884948	AX884948 Sequence
C 573	16.8	80.0	214975	12	AC123335	AC123335 Rattus no	C 646	16.4	78.1	435	2	CQ506365	CQ506365 Sequence
C 574	16.8	80.0	215002	6	AC164614	AC164614 Mus muscu	C 647	16.4	78.1	462	2	CQ411181	CQ411181 Sequence
C 575	16.8	80.0	215073	6	AC157519	AC157519 Mus muscu	C 648	16.4	78.1	472	2	CQ434614	CQ434614 Sequence
C 576	16.8	80.0	215139	12	CR925777	CR925777 Danio rer	C 649	16.4	78.1	472	2	CQ493452	CQ493452 Sequence
C 577	16.8	80.0	215466	12	AC162229	AC162229 Bos tauru	C 650	16.4	78.1	587	2	BD124961	BD124961 Primer fo
C 578	16.8	80.0	215708	11	BX547939	BX547939 Zebrafish	C 651	16.4	78.1	587	2	CQ780252	CQ780252 Sequence
C 579	16.8	80.0	215717	12	AC134484	AC134484 Rattus no	C 652	16.4	78.1	623	7	G92820	G92820 S209P6267FB
C 580	16.8	80.0	215787	12	AC173539	AC173539 Bos tauru	C 653	16.4	78.1	842	2	BD126488	BD126488 Primer fo
C 581	16.8	80.0	217527	12	AC094139	AC094139 Rattus no	C 654	16.4	78.1	842	2	CQ781779	CQ781779 Sequence
C 582	16.8	80.0	219092	12	AC136030	AC136030 Rattus no	C 655	16.4	78.1	964	2	AX771060	AX771060 Sequence
C 583	16.8	80.0	220336	12	AC095220	AC095220 Rattus no	C 656	16.4	78.1	990	2	AR577904	AR577904 Sequence
C 584	16.8	80.0	222772	12	AC095844	AC095844 Rattus no	C 657	16.4	78.1	1076	2	AX400818	AX400818 Sequence
C 585	16.8	80.0	223293	12	AC120325	AC120325 Rattus no	C 658	16.4	78.1	1076	2	AX827362	AX827362 Sequence
C 586	16.8	80.0	223455	12	AC160513	AC160513 Salimtri b	C 659	16.4	78.1	1076	6	AP044574	AP044574 Rattus no
C 587	16.8	80.0	224294	12	AC151861	AC151861 Colobus g	C 660	16.4	78.1	1473	6	BC070959	BC070959 Rattus no
C 588	16.8	80.0	224573	5	AC008758	AC008758 Homo sapi	C 661	16.4	78.1	1919	5	BC106938	BC106938 Homo sapi
C 589	16.8	80.0	227103	12	AC152676	AC152676 Bos tauru	C 662	16.4	78.1	1920	5	BC106937	BC106937 Homo sapi
C 590	16.8	80.0	227343	12	AC133843	AC133843 Rattus no	C 663	16.4	78.1	1972	15	AF457645	AF457645 Listonell
C 591	16.8	80.0	227644	6	AC120347	AC120347 Mus muscu	C 664	16.4	78.1	2476	2	BD127934	BD127934 Primer fo
C 592	16.8	80.0	227695	12	AC115252	AC115252 Rattus no	C 665	16.4	78.1	2476	2	CQ783985	CQ783985 Sequence
C 593	16.8	80.0	228268	12	AC113711	AC113711 Rattus no	C 666	16.4	78.1	2476	5	AK074877	AK074877 Homo sapi
C 594	16.8	80.0	229849	12	AC097285	AC097285 Rattus no	C 667	16.4	78.1	3100	11	AY434090	AY434090 Gallus ga
C 595	16.8	80.0	231245	6	AC159277	AC159277 Mus muscu	C 668	16.4	78.1	3513	14	OCU73126	OCU73126 Oryctolagus
C 596	16.8	80.0	231247	5	AC008543	AC008543 Homo sapi	C 669	16.4	78.1	3555	15	OCU63655	OCU63655 Oryctolagus
C 597	16.8	80.0	231371	12	AC008543	AC008543 Homo sapi	C 670	16.4	78.1	30804	5	AC134338	AC134338 Homo sapi
C 598	16.8	80.0	231371	12	BX284617	BX284617 Homo sapi	C 671	16.4	78.1	34689	6	BX537122	BX537122 Mouse DNA
C 599	16.8	80.0	231371	12	BX284617	BX284617 Homo sapi	C 672	16.4	78.1	35601	5	AC118064	AC118064 Homo sapi
C 600	16.8	80.0	233318	12	AC107514	AC107514 Rattus no	C 673	16.4	78.1	35946	1	AC149043	AC149043 Unculture
C 601	16.8	80.0	234572	12	CT009711	CT009711 Mus muscu	C 674	16.4	78.1	40395	5	AC099491	AC099491 Homo sapi
C 602	16.8	80.0	236382	6	AC133952	AC133952 Mus muscu	C 675	16.4	78.1	40477	12	AC010514	AC010514 Homo sapi

c 676	16.4	78.1	44273	12	AC165330	AC165330 Mus muscu	c 749	16.4	78.1	17570	6	AC111031	AC111031 Mus muscu
c 677	16.4	78.1	50191	5	AL450471	AL450471 Human DNA	c 750	16.4	78.1	179952	6	AC124724	AC124724 Mus muscu
c 678	16.4	78.1	56690	5	AL645761	AL645761 Human DNA	c 751	16.4	78.1	180545	12	CR749169	CR749169 Dantio xer
c 679	16.4	78.1	67272	12	AC100076	AC100076 Mus muscu	c 752	16.4	78.1	180857	12	AC174617	AC174617 Actus nan
c 680	16.4	78.1	76425	5	AC105346	AC105346 Homo sapi	c 753	16.4	78.1	181469	6	AC129222	AC129222 Mus muscu
c 681	16.4	78.1	76702	5	HS756P4	AL035072 Human DNA	c 754	16.4	78.1	182232	6	AC166825	AC166825 Mus muscu
c 682	16.4	78.1	94056	5	AL139192	AL139192 Human DNA	c 755	16.4	78.1	182933	6	AC153972	AC153972 Mus muscu
c 683	16.4	78.1	95682	5	AC104667	AL104667 Homo sapi	c 756	16.4	78.1	183769	12	AC153972	AC153972 Mus muscu
c 684	16.4	78.1	100130	12	AL1391257	AL1391257 Homo sapi	c 757	16.4	78.1	184847	11	AC145510	AC145510 Gasteroat
c 685	16.4	78.1	107107	12	AC150096	AC150096 Gallus ga	c 758	16.4	78.1	184884	12	AC148091	AC148091 Mus muscu
c 686	16.4	78.1	107600	4	ATP16J13	AL049638 Arabidops	c 759	16.4	78.1	185031	6	AC161243	AC161243 Mus muscu
c 687	16.4	78.1	110000	12	AC097542_0	AC097542 Rattus no	c 760	16.4	78.1	186857	6	AL807802	AL807802 Mouse DNA
c 688	16.4	78.1	110000	12	AC155169_4	Continuation (5 of	c 761	16.4	78.1	186857	6	AL807802	AL807802 Mouse DNA
c 689	16.4	78.1	110000	12	BX276116_03	Continuation (4 of	c 762	16.4	78.1	187523	12	AC151181	AC151181 Bos tauru
c 690	16.4	78.1	110000	12	BX276116_04	Continuation (5 of	c 763	16.4	78.1	187639	6	AC161760	AC161760 Mus muscu
c 691	16.4	78.1	110000	11	AB017235_10	Continuation (11 of	c 764	16.4	78.1	187666	6	AC137157	AC137157 Mus muscu
c 692	16.4	78.1	110760	15	AB162345_10	AB162343 Oncorhyn	c 765	16.4	78.1	189186	12	AC15861	AC15861 Mus muscu
c 693	16.4	78.1	113618	12	AC150098	AC150098 Gallus ga	c 766	16.4	78.1	189502	12	AC119987	AC119987 Mus muscu
c 694	16.4	78.1	114800	5	AC123567	AC123567 Homo sapi	c 767	16.4	78.1	190026	4	ATCHR1V33	ATCHR1V33
c 695	16.4	78.1	119882	4	NC586	AL670004 Neurospor	c 768	16.4	78.1	190857	6	AC165317	AC165317 Mus muscu
c 696	16.4	78.1	123872	12	AC159938	AC159938 Taeniopyg	c 769	16.4	78.1	191704	12	AC139737	AC139737 Mus muscu
c 697	16.4	78.1	124758	6	AC111074	AC111074 Mus muscu	c 770	16.4	78.1	191834	12	AC139737	AC139737 Mus muscu
c 698	16.4	78.1	128910	6	AC125308	AC125308 Mus muscu	c 771	16.4	78.1	192134	12	AC174847	AC174847 Mus muscu
c 699	16.4	78.1	132422	6	AC116565	AC116565 Mus muscu	c 772	16.4	78.1	192706	12	AC173393	AC173393 Mus muscu
c 700	16.4	78.1	133485	6	AC118213	AC118213 Mus muscu	c 773	16.4	78.1	192706	12	AC173393	AC173393 Mus muscu
c 701	16.4	78.1	135982	5	HS34606	284487 Human DNA s	c 774	16.4	78.1	194946	12	AC027362	AC027362 Mus muscu
c 702	16.4	78.1	140150	5	AC072057	AC072057 Homo sapi	c 775	16.4	78.1	196421	6	AL672100	AL672100 Mus muscu
c 703	16.4	78.1	141710	12	AC170079	AC170079 Loxodont	c 776	16.4	78.1	196497	6	AC15826	AC15826 Mus muscu
c 704	16.4	78.1	141773	12	AC069348	AC069348 Homo sapi	c 777	16.4	78.1	197360	12	AC018829	AC018829 Homo sapi
c 705	16.4	78.1	142982	5	CR376760	CR376760 Pan trogl	c 778	16.4	78.1	197574	6	AC118704	AC118704 Mus muscu
c 706	16.4	78.1	143413	6	AC1025979	AC1025979 Homo sapi	c 779	16.4	78.1	197782	5	AC012377	AC012377 Mus muscu
c 707	16.4	78.1	145213	12	AC154402	AC154402 Mus muscu	c 780	16.4	78.1	198647	6	AL662868	AL662868 Mus muscu
c 708	16.4	78.1	145435	12	AC026685	AC026685 Homo sapi	c 781	16.4	78.1	198647	12	AC096941	AC096941 Mus muscu
c 709	16.4	78.1	149105	6	AC127263	AC127263 Mus muscu	c 782	16.4	78.1	200542	5	CNS01RG3	CNS01RG3 Mus muscu
c 710	16.4	78.1	149235	12	AC068266	AC068266 Homo sapi	c 783	16.4	78.1	201259	6	AC159504	AC159504 Mus muscu
c 711	16.4	78.1	149450	6	AC125218	AC125218 Mus muscu	c 784	16.4	78.1	201728	6	AC159504	AC159504 Mus muscu
c 712	16.4	78.1	151649	12	AC179314	AC179314 Strongylo	c 785	16.4	78.1	204000	6	AC136645	AC136645 Mus muscu
c 713	16.4	78.1	152846	5	AC025709	AC025709 Homo sapi	c 786	16.4	78.1	204354	6	AC121795	AC121795 Mus muscu
c 714	16.4	78.1	155127	12	AC008532	AC008532 Homo sapi	c 787	16.4	78.1	207019	12	AC140332	AC140332 Mus muscu
c 715	16.4	78.1	155313	12	AC068315	AC068315 Homo sapi	c 788	16.4	78.1	207149	6	AC140332	AC140332 Mus muscu
c 716	16.4	78.1	156632	12	AC155063	AC155063 Bos tauru	c 789	16.4	78.1	208096	6	AC122242	AC122242 Mus muscu
c 717	16.4	78.1	158218	6	CT009544	CT009544 Mouse DNA	c 790	16.4	78.1	209934	6	AC154216	AC154216 Mus muscu
c 718	16.4	78.1	158469	6	AC121796	AC121796 Mus muscu	c 791	16.4	78.1	215909	12	AC173014	AC173014 Bos tauru
c 719	16.4	78.1	158620	6	AC121796	AC121796 Mus muscu	c 792	16.4	78.1	216721	6	AC154200	AC154200 Mus muscu
c 720	16.4	78.1	158709	12	AC083827	AC083827 Homo sapi	c 793	16.4	78.1	216911	12	AC018809	AC018809 Homo sapi
c 721	16.4	78.1	158838	5	CNS05TDN	AL356805 Human chr	c 794	16.4	78.1	218213	12	AC109161	AC109161 Mus muscu
c 722	16.4	78.1	158901	5	AC093852	AC093852 Homo sapi	c 795	16.4	78.1	222259	12	AC110858	AC110858 Rattus no
c 723	16.4	78.1	159260	12	AC020789	AC020789 Mus muscu	c 796	16.4	78.1	222753	12	AC150949	AC150949 Bos tauru
c 724	16.4	78.1	161903	12	AC021996	AC021996 Homo sapi	c 797	16.4	78.1	223469	12	AC112451	AC112451 Mus muscu
c 725	16.4	78.1	161955	6	AC149598	AC149598 Mus muscu	c 798	16.4	78.1	224742	12	AC150075	AC150075 Mus muscu
c 726	16.4	78.1	162552	12	AC162862	AC162862 Mus muscu	c 799	16.4	78.1	225020	6	AC109247	AC109247 Mus muscu
c 727	16.4	78.1	162808	12	AC021340	AC021340 Homo sapi	c 800	16.4	78.1	226843	12	AC157187	AC157187 Bos tauru
c 728	16.4	78.1	163235	12	AC158934	AC158934 Bos tauru	c 801	16.4	78.1	227029	12	AC125976	AC125976 Mus muscu
c 729	16.4	78.1	163714	12	AC150113	AC150113 Gallus ga	c 802	16.4	78.1	227047	12	AC156462	AC156462 Bos tauru
c 730	16.4	78.1	163897	6	AC132256	AC132256 Mus muscu	c 803	16.4	78.1	227759	12	AC107446	AC107446 Mus muscu
c 731	16.4	78.1	164297	12	AC012138	AC012138 Homo sapi	c 804	16.4	78.1	229171	6	CT009762	CT009762 Mouse DNA
c 732	16.4	78.1	164739	5	BX682546	BX682546 Homo sapi	c 805	16.4	78.1	229478	12	AC151871	AC151871 Lemur cat
c 733	16.4	78.1	165639	12	AL513023	AL513023 Human DNA	c 806	16.4	78.1	229940	12	AC097414	AC097414 Rattus no
c 734	16.4	78.1	166110	11	AL929305	AL929305 Zebrafish	c 807	16.4	78.1	230352	12	AC141213	AC141213 Rattus no
c 735	16.4	78.1	166474	12	AC162362	AC162362 Mus muscu	c 808	16.4	78.1	230352	12	AC168897	AC168897 Bos tauru
c 736	16.4	78.1	166534	6	AC155295	AC155295 Mus muscu	c 809	16.4	78.1	232346	6	CR974489	CR974489 Mouse DNA
c 737	16.4	78.1	166871	6	AC120381	AC120381 Mus muscu	c 810	16.4	78.1	232866	12	AC174949	AC174949 Bos tauru
c 738	16.4	78.1	167390	5	AC007263	AC007263 Homo sapi	c 811	16.4	78.1	234316	12	AC098553	AC098553 Rattus no
c 739	16.4	78.1	171916	12	AC104909	AC104909 Mus muscu	c 812	16.4	78.1	235102	12	AC118993	AC118993 Rattus no
c 740	16.4	78.1	173145	5	AC026269	AC026269 Homo sapi	c 813	16.4	78.1	235918	12	AC129279	AC129279 Mus muscu
c 741	16.4	78.1	173336	5	AC025168	AC025168 Homo sapi	c 814	16.4	78.1	236372	12	AC098803	AC098803 Rattus no
c 742	16.4	78.1	174773	11	CR383662	CR383662 Zebrafish	c 815	16.4	78.1	238312	12	AC152282	AC152282 Bos tauru
c 743	16.4	78.1	175494	5	AL590667	AL590667 Human DNA	c 816	16.4	78.1	239100	12	AC105455	AC105455 Rattus no
c 744	16.4	78.1	175833	12	AC073412	AC073412 Homo sapi	c 817	16.4	78.1	241006	12	AC152096	AC152096 Bos tauru
c 745	16.4	78.1	175864	12	AC133421	AC133421 Rattus no	c 818	16.4	78.1	241126	12	AC095541	AC095541 Rattus no
c 746	16.4	78.1	176056	5	AC073437	AC073437 Mus muscu	c 819	16.4	78.1	244242	12	AC107551	AC107551 Rattus no
c 747	16.4	78.1	176421	5	AC096888	AC096888 Homo sapi	c 820	16.4	78.1	244478	12	AC130167	AC130167 Rattus no
c 748	16.4	78.1	177344	6	AC140349	AC140349 Mus muscu	c 821	16.4	78.1	245809	12	AC105713	AC105713 Rattus no

C 822	16.4	78.1	245907	6	AC097752	AC097752 Rattus no	C 895	16.2	77.1	1568	2	CS200399	CS200399 Sequence
C 823	16.4	78.1	246101	12	AC135646	AC135646 Rattus no	C 896	16.2	77.1	1632	6	BC079298	BC079298 Rattus no
C 824	16.4	78.1	246716	12	AC125644	AC125644 Rattus no	C 897	16.2	77.1	1790	2	BD224541	BD224541 Materials
C 825	16.4	78.1	248381	12	AC009937	AC009937 Homo sapi	C 898	16.2	77.1	1790	2	AR216591	AR216591 Sequence
C 826	16.4	78.1	248825	12	AC131555	AC131555 Rattus no	C 899	16.2	77.1	1809	4	AY163489	AY163489 Nicotiana
C 827	16.4	78.1	249622	6	AC124777	AC124777 Mus muscu	C 900	16.2	77.1	1881	4	CQ732800	CQ732800 Sequence
C 828	16.4	78.1	249622	12	AC132300	AC132300 Rattus no	C 901	16.2	77.1	2046	2	CQ722632	CQ722632 Sequence
C 829	16.4	78.1	250785	12	AC137355	AC137355 Rattus no	C 902	16.2	77.1	2238	6	BC062614	BC062614 Homo sapi
C 830	16.4	78.1	253582	12	AC174812	AC174812 Rattus no	C 903	16.2	77.1	2239	6	AF085219	AF085219 Mus muscu
C 831	16.4	78.1	256796	12	AC094838	AC094838 Rattus no	C 904	16.2	77.1	2240	5	BC035288	BC035288 Homo sapi
C 832	16.4	78.1	258368	12	AC110369	AC110369 Rattus no	C 905	16.2	77.1	2344	5	BC050578	BC050578 Homo sapi
C 833	16.4	78.1	259679	12	AC169684	AC169684 Bos tauri	C 906	16.2	77.1	2438	4	AB213565	AB213565 Vitis vin
C 834	16.4	78.1	262289	12	AC117951	AC117951 Rattus no	C 907	16.2	77.1	2447	2	AX713869	AX713869 Sequence
C 835	16.4	78.1	264075	12	AC134075	AC134075 Rattus no	C 908	16.2	77.1	2447	5	AK055758	AK055758 Homo sapi
C 836	16.4	78.1	264623	12	AC124767	AC124767 Mus muscu	C 909	16.2	77.1	2538	2	CS167823	CS167823 Sequence
C 837	16.4	78.1	270878	12	AC114433	AC114433 Rattus no	C 910	16.2	77.1	2538	2	AX833944	AX833944 Sequence
C 838	16.4	78.1	275383	12	AC096227	AC096227 Rattus no	C 911	16.2	77.1	2538	5	AK096080	AK096080 Homo sapi
C 839	16.4	78.1	281043	12	AC119764	AC119764 Rattus no	C 912	16.2	77.1	2811	2	AX961936	AX961936 Sequence
C 840	16.4	78.1	297398	12	AC097578	AC097578 Rattus no	C 913	16.2	77.1	2842	6	BC006713	BC006713 Mus muscu
C 841	16.2	77.1	23	2	AR709482	AR709482 Sequence	C 914	16.2	77.1	2872	4	AY163490	AY163490 Nicotiana
C 842	16.2	77.1	31	2	BD225203	BD225203 Human pap	C 915	16.2	77.1	2966	5	AK095962	AK095962 Homo sapi
C 843	16.2	77.1	31	2	BD225204	BD225204 Human pap	C 916	16.2	77.1	2968	11	AF197138	AF197138 Fugu rubr
C 844	16.2	77.1	31	2	CS195919	CS195919 Sequence	C 917	16.2	77.1	3043	2	CS164047	CS164047 Sequence
C 845	16.2	77.1	31	2	CS195920	CS195920 Sequence	C 918	16.2	77.1	3043	2	CS164691	CS164691 Sequence
C 846	16.2	77.1	120	5	AF313103	AF313103 Homo sapi	C 919	16.2	77.1	3043	6	S49760	S49760 diacylglyce
C 847	16.2	77.1	274	6	MMU5489	MMU5489 Mus muscu	C 920	16.2	77.1	3058	2	CO851151	CO851151 Sequence
C 848	16.2	77.1	401	2	CO524563	CO524563 Sequence	C 921	16.2	77.1	3058	5	AK095962	AK095962 Homo sapi
C 849	16.2	77.1	447	10	DQ057230	DQ057230 Human pap	C 922	16.2	77.1	3100	5	AB220457	AB220457 Macaca fa
C 850	16.2	77.1	447	10	DQ057291	DQ057291 Human pap	C 923	16.2	77.1	3115	5	AK123589	AK123589 Homo sapi
C 851	16.2	77.1	447	10	DQ057292	DQ057292 Human pap	C 924	16.2	77.1	3174	2	CS168909	CS168909 Sequence
C 852	16.2	77.1	447	10	DQ057293	DQ057293 Human pap	C 925	16.2	77.1	3174	2	AX835030	AX835030 Sequence
C 853	16.2	77.1	447	10	DQ057294	DQ057294 Human pap	C 926	16.2	77.1	3174	5	AK093792	AK093792 Homo sapi
C 854	16.2	77.1	450	10	DQ057309	DQ057309 Human pap	C 927	16.2	77.1	3192	2	E02295	E02295 Xeni-Puvit
C 855	16.2	77.1	450	10	DQ057310	DQ057310 Human pap	C 928	16.2	77.1	3199	11	BC110725	BC110725 Xenopus 1
C 856	16.2	77.1	450	10	DQ057311	DQ057311 Human pap	C 929	16.2	77.1	3490	6	RAT17PBCA	RAT17PBCA Xenopus 1
C 857	16.2	77.1	450	10	DQ057312	DQ057312 Human pap	C 930	16.2	77.1	3656	11	BC042353	BC042353 Xenopus 1
C 858	16.2	77.1	450	10	DQ057313	DQ057313 Human pap	C 931	16.2	77.1	6327	5	HSM804618	HSM804618 Homo sapi
C 859	16.2	77.1	450	10	DQ057314	DQ057314 Human pap	C 932	16.2	77.1	7851	10	PH3350G	PH3350G Homo sapi
C 860	16.2	77.1	596	7	G38228	G38228 RCT-5-841D	C 933	16.2	77.1	7942	10	HPV52	HPV52 Homo sapi
C 861	16.2	77.1	609	7	BV414049	BV414049 S229P6495	C 934	16.2	77.1	7942	10	HPV52	HPV52 Homo sapi
C 862	16.2	77.1	622	7	BV563272	BV563272 qu984E05	C 935	16.2	77.1	11709	12	AC114711	AC114711 ⁴
C 863	16.2	77.1	632	7	G88736	G88736 S209P6370FH	C 936	16.2	77.1	15580	12	AC013045	AC013045 Drosophi
C 864	16.2	77.1	632	7	BV324835	BV324835 S241P6168	C 937	16.2	77.1	16370	11	AF197897	AF197897 Takifugu
C 865	16.2	77.1	680	7	BV579495	BV579495 G591P6132	C 938	16.2	77.1	23992	5	AC132004	AC132004 Homo sapi
C 866	16.2	77.1	684	7	BV039482	BV039482 S212P6056	C 939	16.2	77.1	25738	5	AF125183	AF125183 Homo sapi
C 867	16.2	77.1	697	7	BV567437	BV567437 qc701D04	C 940	16.2	77.1	26995	5	AC110490	AC110490 Homo sapi
C 868	16.2	77.1	699	7	BV255124	BV255124 S234P6182	C 941	16.2	77.1	31632	5	AC104520	AC104520 Homo sapi
C 869	16.2	77.1	728	7	BV662556	BV662556 S216P6117	C 942	16.2	77.1	31632	5	HSO19A	HSO19A Homo sapi
C 870	16.2	77.1	752	7	BV642966	BV642966 S216P6355	C 943	16.2	77.1	40558	5	AF087017	AF087017 Homo sapi
C 871	16.2	77.1	795	2	BD124746	BD124746 Primer fo	C 944	16.2	77.1	41407	5	AC053467	AC053467 Homo sapi
C 872	16.2	77.1	795	2	BD126759	BD126759 Primer fo	C 945	16.2	77.1	41605	6	AL928923	AL928923 Mouse DNA
C 873	16.2	77.1	795	2	CQ780037	CQ780037 Sequence	C 946	16.2	77.1	41770	5	AC127090	AC127090 Homo sapi
C 874	16.2	77.1	795	2	CQ780037	CQ780037 Sequence	C 947	16.2	77.1	44924	5	AC101609	AC101609 Homo sapi
C 875	16.2	77.1	803	10	PPHE67	PPHE67 Sequence	C 948	16.2	77.1	45746	5	AC005619	AC005619 Homo sapi
C 876	16.2	77.1	811	7	BV623241	BV623241 Human pap1	C 949	16.2	77.1	46638	5	AC104823	AC104823 Homo sapi
C 877	16.2	77.1	816	7	BV590311	BV590311 G591P6426	C 950	16.2	77.1	55641	6	CR751603	CR751603 Mouse DNA
C 878	16.2	77.1	823	7	BV625404	BV625404 S217P6793	C 951	16.2	77.1	57835	12	AC135327	AC135327 Homo sapi
C 879	16.2	77.1	835	7	BV538314	BV538314 G591P6385	C 952	16.2	77.1	60773	12	AC069576	AC069576 Homo sapi
C 880	16.2	77.1	848	7	BV043791	BV043791 S212P6170	C 953	16.2	77.1	63509	12	AC176482	AC176482 Strongylo
C 881	16.2	77.1	855	7	BV503424	BV503424 QK196A07	C 954	16.2	77.1	63736	5	AC127453	AC127453 Homo sapi
C 882	16.2	77.1	866	7	BV584846	BV584846 G591P6275	C 955	16.2	77.1	64200	5	AC087691	AC087691 Homo sapi
C 883	16.2	77.1	930	2	BD192646	BD192646 123 human	C 956	16.2	77.1	65609	5	AC025858	AC025858 Homo sapi
C 884	16.2	77.1	930	2	BD192646	BD192646 123 human	C 957	16.2	77.1	65750	5	AC005003	AC005003 Homo sapi
C 885	16.2	77.1	930	2	CS129731	CS129731 Sequence	C 958	16.2	77.1	65960	12	AC079627	AC079627 Mus muscu
C 886	16.2	77.1	930	2	CS130002	CS130002 Sequence	C 959	16.2	77.1	67911	5	AP001413	AP001413 Homo sapi
C 887	16.2	77.1	930	2	AR184117	AR184117 Sequence	C 960	16.2	77.1	68166	12	AC080472	AC080472 Homo sapi
C 888	16.2	77.1	930	2	AR184202	AR184202 Sequence	C 961	16.2	77.1	68227	14	AC144754	AC144754 Onitichor
C 889	16.2	77.1	930	2	AR703286	AR703286 Sequence	C 962	16.2	77.1	70148	5	AL357061	AL357061 Human DNA
C 890	16.2	77.1	930	2	AR703237	AR703237 Sequence	C 963	16.2	77.1	70313	12	AR659673	AR659673 Sequence
C 891	16.2	77.1	930	2	AR703237	AR703237 Sequence	C 964	16.2	77.1	70313	12	AR659673	AR659673 Sequence
C 892	16.2	77.1	1058	4	IMES57240	IMES57240 Ilex x me	C 965	16.2	77.1	71174	5	AL5590729	AL5590729 Human DNA
C 893	16.2	77.1	1342	4	AY355917	AY355917 Strombomo	C 966	16.2	77.1	72347	2	AR655952	AR655952 Sequence
C 894	16.2	77.1	1524	4	AF483211	AF483211 Alonboa m	C 967	16.2	77.1	75164	5	AC093867	AC093867 Homo sapi

```

c 968 16.2 77.1 76228 12 AC016338
c 969 16.2 77.1 76540 5 AC003678
c 970 16.2 77.1 77313 5 AC110757
c 971 16.2 77.1 77573 5 AY572796
c 972 16.2 77.1 79594 5 AL138805
c 973 16.2 77.1 81542 5 AC100830
c 974 16.2 77.1 82010 5 BS000122
c 975 16.2 77.1 82514 5 AC004137
c 976 16.2 77.1 86480 5 CR759904
c 977 16.2 77.1 86483 12 CR759868
c 978 16.2 77.1 86576 5 AL929403
c 979 16.2 77.1 87364 12 AC138702
c 980 16.2 77.1 88484 5 AL355794
c 981 16.2 77.1 89177 5 AL356000
c 982 16.2 77.1 89214 5 AC093009
c 983 16.2 77.1 92134 12 AC018406
c 984 16.2 77.1 92487 6 BX511247
c 985 16.2 77.1 92644 12 AF186995
c 986 16.2 77.1 93292 5 CR759829
c 987 16.2 77.1 93732 12 AC136444
c 988 16.2 77.1 95930 5 AC106033
c 989 16.2 77.1 96037 12 AC171163
c 990 16.2 77.1 96120 5 AC093669
c 991 16.2 77.1 96798 5 AL356420
c 992 16.2 77.1 96877 12 AC017776
c 993 16.2 77.1 97466 5 BX005422
c 994 16.2 77.1 98835 5 AC073195
c 995 16.2 77.1 99088 5 AC093401
c 996 16.2 77.1 100000 5 AP000092
c 997 16.2 77.1 100000 5 AP000153
c 998 16.2 77.1 100000 5 AP000196
c 999 16.2 77.1 100293 5 AP006259
1000 16.2 77.1 100316 5 AC004556

```

ALIGNMENTS

```

RESULT 1
AR349145 21 bp DNA linear PAT 17-AUG-2003
LOCUS AR349145 Sequence 85 from patent US 6583278.
ACCESSION AR349145
VERSION AR349145.1 GI:33749850
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 21)
AUTHORS Carter N.M.
TITLE Nucleic acid probes complementary to human papilloma virus nucleic acid
JOURNAL Patent: US 6583278-A 85 24-JUN-2003;
Gen-Probe Incorporated; San Diego, CA
FEATURES
source Location/Qualifiers
1..21
/mol_type="genomic DNA"
ORIGIN
Query Match 100.0%; Score 21; DB 2; Length 21;
Best Local Similarity 100.0%; Pred. No. 0.26;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 CAGGACACAGTGCGCTTTTGAC 21
Db 1 CAGGACACAGTGCGCTTTTGAC 21
RESULT 2
AR349146 21 bp DNA linear PAT 17-AUG-2003
LOCUS AR349146/c
DEFINITION Sequence 86 from patent US 6583278.

```

```

ACCESSION AR349146
VERSION AR349146.1 GI:33749851
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 21)
AUTHORS Carter N.M.
TITLE Nucleic acid probes complementary to human papilloma virus nucleic acid
JOURNAL Patent: US 6583278-A 86 24-JUN-2003;
Gen-Probe Incorporated; San Diego, CA
FEATURES
source Location/Qualifiers
1..21
/mol_type="genomic DNA"
ORIGIN
Query Match 100.0%; Score 21; DB 2; Length 21;
Best Local Similarity 100.0%; Pred. No. 0.26;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 CAGGACACAGTGCGCTTTTGAC 21
Db 21 CAGGACACAGTGCGCTTTTGAC 1
RESULT 3
AR349147 21 bp DNA linear PAT 17-AUG-2003
LOCUS AR349147 Sequence 87 from patent US 6583278.
ACCESSION AR349147
VERSION AR349147.1 GI:33749852
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 21)
AUTHORS Carter N.M.
TITLE Nucleic acid probes complementary to human papilloma virus nucleic acid
JOURNAL Patent: US 6583278-A 87 24-JUN-2003;
Gen-Probe Incorporated; San Diego, CA
FEATURES
source Location/Qualifiers
1..21
/mol_type="genomic DNA"
ORIGIN
Query Match 100.0%; Score 21; DB 2; Length 21;
Best Local Similarity 100.0%; Pred. No. 0.26;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 CAGGACACAGTGCGCTTTTGAC 21
Db 1 CAGGACACAGTGCGCTTTTGAC 21
RESULT 4
AR349148 21 bp DNA linear PAT 17-AUG-2003
LOCUS AR349148/c
DEFINITION Sequence 88 from patent US 6583278.
ACCESSION AR349148
VERSION AR349148.1 GI:33749853
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 21)
AUTHORS Carter N.M.
TITLE Nucleic acid probes complementary to human papilloma virus nucleic acid
JOURNAL Patent: US 6583278-A 88 24-JUN-2003;

```


Gen-Probe Incorporated; San Diego, CA

FEATURES
source
1. .21
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 100.0%; Score 21; DB 2; Length 21;
Best Local Similarity 100.0%; Pred. No. 0.26;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY
1 CAGGACACAGTGGCTTTGAC 21
|||||
21 CAGGACACAGTGGCTTTGAC 1

Db
1 CAGGACACAGTGGCTTTGAC 1

RESULT 5
DD187316 104 bp DNA linear PAT 19-JAN-2006
DEFINITION SPICEOSOME MEDIATED RNA TRANS-SPLICING.
ACCESSION DD187316 GI:85643668
KEYWORDS JP 2005176849-A/118.
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.

REFERENCE
AUTHORS Walsh, C., Chao, H., Mansfield, G.S., Pataraju, M., Baker, C.C.,
GarciaBlanco, M.A. and Mitchell, L.G.
TITLE SPICEOSOME MEDIATED RNA TRANS-SPLICING
JOURNAL Patent: JP 2005176849-A 118 07-JUL-2005;
INTRON INC

COMMENT
OS Artificial Sequence
PN JP 2005176849-A/118
PD 07-JUL-2005
PR 07-JAN-2005 JP 2005003076
PR 08-JAN-2001 US 09/756095, 08-JAN-2001 US 09/756096, PR
08-JAN-2001 US 09/756097, 20-APR-2001 US 09/838858, PR
29-AUG-2001 US 09/941492
PI christopher walsh, hengjun chao, gary s mansfield, madaia pi
ptaraju,
PI carl c baker, mariano a garcia-blanco, lloyd g mitchell CC
Binding domain of human papilloma virus PTM
FH Key Location/Qualifiers

FEATURES
source
1. .104
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"

ORIGIN
Query Match 100.0%; Score 21; DB 2; Length 104;
Best Local Similarity 100.0%; Pred. No. 0.28;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY
1 CAGGACACAGTGGCTTTGAC 21
|||||
4 CAGGACACAGTGGCTTTGAC 24

Db
4 CAGGACACAGTGGCTTTGAC 24

RESULT 6
DD187319 138 bp DNA linear PAT 19-JAN-2006
DEFINITION SPICEOSOME MEDIATED RNA TRANS-SPLICING.
ACCESSION DD187319 GI:85643671
KEYWORDS JP 2005176849-A/121.
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.

REFERENCE
AUTHORS Walsh, C., Chao, H., Mansfield, G.S., Pataraju, M., Baker, C.C.,
GarciaBlanco, M.A. and Mitchell, L.G.

SPICEOSOME MEDIATED RNA TRANS-SPLICING
Patent: JP 2005176849-A 121 07-JUL-2005;
INTRON INC

COMMENT
OS Artificial Sequence
PN JP 2005176849-A/121
PD 07-JUL-2005
PR 07-JAN-2005 JP 2005003076
PR 08-JAN-2001 US 09/756095, 08-JAN-2001 US 09/756096, PR
08-JAN-2001 US 09/756097, 20-APR-2001 US 09/838858, PR
29-AUG-2001 US 09/941492
PI christopher walsh, hengjun chao, gary s mansfield, madaia pi
ptaraju,
PI carl c baker, mariano a garcia-blanco, lloyd g mitchell CC
Binding domain of human papilloma virus PTM
FH Key Location/Qualifiers

FEATURES
source
1. .138
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"

ORIGIN
Query Match 100.0%; Score 21; DB 2; Length 138;
Best Local Similarity 100.0%; Pred. No. 0.29;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY
1 CAGGACACAGTGGCTTTGAC 21
|||||
39 CAGGACACAGTGGCTTTGAC 59

Db
39 CAGGACACAGTGGCTTTGAC 59

RESULT 7
HPUS9900/c 170 bp mRNA linear VRL 06-JUN-1996
LOCUS HPUS9900
DEFINITION Human papillomavirus type 16 mutant442, major E6E7 splice variant
mRNA, partial cds.
ACCESSION US9900
VERSION US9900.1 GI:1405361
KEYWORDS Human papillomavirus type 16
Human papillomavirus type 16
Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
Alpha papillomavirus.
1 (bases 1 to 170)

REFERENCE
AUTHORS Shippey, R., Siwkowski, A. and Hampel, A.
TITLE Direct Submission
JOURNAL Submitted (03-JUN-1996) Biology, N.I.U., Montgomery Hall #319,
DeKalb, IL 60115, USA

FEATURES
source
1. .170
/organism="Human papillomavirus type 16"
/mol_type="mRNA"
/strain="SiHA"
/db_xref="taxon:333760"
/map="155-506 in the genomic map"
/note="The mutation at site 442 in the E6E7 major splice
variant was verified with a series of clones and PCR batch
sequencing"
1. .81
/note="truncated E6 protein product of the major splice
variant"
/codon_start=1
/product="E6E7"
/protein_id="AA803505.1"
/db_xref="GI:1405362"
/translation="ELQVTIHDIILCYCKQQLRREYV"
1. .72
73. .>170
105
/note="site of mutation (nucleotide 442 in genome); This
mutation results in an amino acid change from B to D when
the full-size transcript is translated"

ORIGIN
exon
exon
misc_feature

Query Match 100.0%; Score 21; DB 10; Length 170;
Best Local Similarity 100.0%; Pred. No. 0.29;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAGGACACAGTGGCTTTGAC 21
|||||
Db 104 CAGGACACAGTGGCTTTGAC 84

RESULT 8
HPV14515/c 245 bp DNA linear VRL 15-FEB-2001
LOCUS
DEFINITION Human papillomavirus clone HPV1629 B6 protein (B6) gene, partial
cds.
ACCESSION U14515
VERSION U14515.1 GI:984951
KEYWORDS
SOURCE Human papillomavirus
ORGANISM Viruses; deDNA viruses, no RNA stage; Papillomaviridae;
unclassified Papillomaviridae.

REFERENCE 1 (bases 1 to 245)
AUTHORS Haeger, D.G., Galutira, D.F. and Younghusband, B.H.
JOURNAL Sequence variation in the B6 gene of human papillomavirus type 16
REFERENCE 2 (bases 1 to 245)
AUTHORS Galutira, D.F.
TITLE Direct Submision
SUBMITTED (07-SEP-1994) Dante F. Galutira, Memorial University of
Newfoundland, Discipline of Pathology, Fac. Medicine, St. John's,
Newfoundland, A1B 3V6, Canada
JOURNAL Location/Qualifiers

FEATURES
source
1..245
/organism="Human papillomavirus"
/mol_type="genomic DNA"
/strain="HPV16"
/db_xref="taxon:10566"
/clone="HPV1629"
/tissue_type="cervical tissue from invasive carcinoma"

gene
1..245
/gene="B6"
/note="T to A transversion from PA16, GenBank Accession
Number K02718, position 286; NCBI gi: 333031"
/replace="t"
64
/gene="B6"
/note="A to G transversion from PA16, GenBank Accession
Number K02718, position 289; NCBI gi: 333031"
/replace="a"
110
/gene="B6"
/note="C to T transition from PA16, GenBank Accession
Number K02718, position 335; NCBI gi: 333031"
/replace="c"
125
/gene="B6"
/note="T to G transversion from PA16, GenBank Accession
Number K02718, position 350; NCBI gi: 333031"
/replace="t"
139
/gene="B6"
/note="A to T transversion from PA16, GenBank Accession

ORIGIN
Number K02718, position 364; NCBI gi: 333031"
/replace="a"

Query Match 100.0%; Score 21; DB 10; Length 245;
Best Local Similarity 100.0%; Pred. No. 0.29;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAGGACACAGTGGCTTTGAC 21
|||||
Db 215 CAGGACACAGTGGCTTTGAC 195

RESULT 9
HPV14511/c 253 bp DNA linear VRL 15-FEB-2001
LOCUS
DEFINITION Human papillomavirus clone HPV1601 B6 protein (B6) gene, partial
cds.
ACCESSION U14511
VERSION U14511.1 GI:984943
KEYWORDS
SOURCE Human papillomavirus
ORGANISM Viruses; deDNA viruses, no RNA stage; Papillomaviridae;
unclassified Papillomaviridae.

REFERENCE 1 (bases 1 to 253)
AUTHORS Haeger, D.G., Galutira, D.F. and Younghusband, B.H.
JOURNAL Sequence variation in the B6 gene of human papillomavirus type 16
REFERENCE 2 (bases 1 to 253)
AUTHORS Galutira, D.F.
TITLE Direct Submision
SUBMITTED (07-SEP-1994) Dante F. Galutira, Memorial University of
Newfoundland, Discipline of Pathology, Fac. Medicine, St. John's,
Newfoundland, A1B 3V6, Canada
JOURNAL Location/Qualifiers

FEATURES
source
1..253
/organism="Human papillomavirus"
/mol_type="genomic DNA"
/strain="HPV16"
/db_xref="taxon:10566"
/clone="HPV1601"
/cell_line="Caaki and Sila"
/tissue_type="cervical"

gene
1..253
/gene="B6"
/note="T to G transversion from PA16, GenBank Accession
Number K02718, position 350; NCBI gi: 333031"
/replace="t"
115
/gene="B6"
/note="A to G transversion from PA16, GenBank Accession
Number K02718, position 289; NCBI gi: 333031"
/replace="a"
125
/gene="B6"
/note="T to G transversion from PA16, GenBank Accession
Number K02718, position 335; NCBI gi: 333031"
/replace="c"
139
/gene="B6"
/note="A to T transversion from PA16, GenBank Accession

variation
/gene="B6"
/note="T to A transversion from PA16, GenBank Accession
Number K02718, position 286; NCBI gi: 333031"
/replace="t"
64
/gene="B6"
/note="A to G transversion from PA16, GenBank Accession
Number K02718, position 289; NCBI gi: 333031"
/replace="a"
110
/gene="B6"
/note="C to T transition from PA16, GenBank Accession
Number K02718, position 335; NCBI gi: 333031"
/replace="c"
125
/gene="B6"
/note="T to G transversion from PA16, GenBank Accession
Number K02718, position 350; NCBI gi: 333031"
/replace="t"
139
/gene="B6"
/note="A to T transversion from PA16, GenBank Accession

ORIGIN
Number K02718, position 364; NCBI gi: 333031"
/replace="a"

Query Match 100.0%; Score 21; DB 10; Length 253;
Best Local Similarity 100.0%; Pred. No. 0.29;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAGGACACAGTGGCTTTGAC 21
|||||
Db 205 CAGGACACAGTGGCTTTGAC 185

RESULT 10
HPV14512/c 271 bp DNA linear VRL 14-SEP-1995

DEFINITION	Accession	Version	Keywords	Source	Organism
Human papillomavirus clone HPV1603 E6 protein (E6) gene, partial cds.	U14512	1	GI:984945		
Human papillomavirus					
Human papillomavirus					
Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;					
Unclassified Papillomaviridae.					
1 (bases 1 to 271)					
Haegert,D.G., Galutira,D.F. and Youngusband,B.H.					
Sequence variation in the E6 gene of human papillomavirus type 16					
Unpublished					
2 (bases 1 to 271)					
Galutira,D.F.					
Direct Submission					
Submitted (07-SEP-1994) Dante F. Galutira, Memorial University of					
Newfoundland, Discipline of Pathology, Fac. Medicine, St. John's,					
Newfoundland, A1B 3X6, Canada					
Location/Qualifiers					
1..271					
/organism="Human papillomavirus"					
/mol_type="genomic DNA"					
/strain="HPV16"					
/db_xref="taxon:10566"					
/clone="HPV1603"					
/tissue_type="cervical tissue from invasive carcinoma"					
1..271					
/gene="E6"					
<1..>271					
/gene="E6"					
/codon_start=2					
/product="E6 protein"					
/protein_id="AAB60566.1"					
/db_xref="GI:984946"					
/translation="VYDPAFRLDLYRDNQVPAVCDCKLGYEKISEYRHYCYSLYG					
TTLEQNTKPLCDLLIRCTNCQKPLCPBEKORHLDKORFHNIRG"					
85					
/gene="E6"					
/note="T to G transversion from PA16, GenBank Accession					
Number K02718, position 310; NCBI gi: 333031"					
/replace="T"					
variation					
Query Match					
Best Local Similarity					
100.0%; Score 21; DB 10; Length 271;					
100.0%; Pred. No. 0.3;					
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;					
ORIGIN					
Query					
1 CAGACACAGTGGCTTTTGAC 21					
Db					
215 CAGACACAGTGGCTTTTGAC 195					
RESULT 11					
HPV14513/c					
LOCUS					
DEFINITION					
Human papillomavirus clone HPV1607 E6 protein (E6) gene, partial					
cds.					
Accession					
U14513					
Version					
U14513.1					
Keywords					
GI:984947					
Source					
Human papillomavirus					
Human papillomavirus					
Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;					
Unclassified Papillomaviridae.					
1 (bases 1 to 272)					
Haegert,D.G., Galutira,D.F. and Youngusband,B.H.					
Sequence variation in the E6 gene of human papillomavirus type 16					
Unpublished					
2 (bases 1 to 272)					
Galutira,D.F.					
Direct Submission					
Submitted (07-SEP-1994) Dante F. Galutira, Memorial University of					

FEATURES							
source							
Newfoundland, Discipline of Pathology, Fac. Medicine, St. John's,							
Newfoundland, A1B 3V6, Canada							
Location/Qualifiers							
1..272							
/organism="Human papillomavirus"							
/mol_type="genomic DNA"							
/strain="HPV16"							
/db_xref="taxon:10566"							
/clone="HPV1607"							
/tissue_type="cervical tissue from invasive carcinoma"							
1..272							
/gene="E6"							
<1..>272							
/gene="E6"							
/codon_start=2							
/product="E6 protein"							
/protein_id="AAB60567.2"							
/translation="VFDFPRDMMYTRDGNPPAVCDCKLKFYSKISEFYHYCYSLYG TTLBQGNKPLCDLLIRNCINCPKPCPEEKORHLDRKORPHNIGR" 31							
/gene="E6"							
/note="C to G transversion from PA16, GenBank Accession Number K02718, position 256; NCBI gi: 333031"							
/replace="C"							
ORIGIN							
Query Match 100.0%; Score 21; DB 10; Length 272;							
Best Local Similarity 100.0%; Pred. No. 0.3;							
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;							
Cy 1 CAGGACACAGTGGCTTTTGAC 21							
Db 215 CAGGACACAGTGGCTTTTGAC 195							
RESULT 12							
HPV14514 273 bp DNA linear VRL 15-FEB-2001							
LOCUS HPV14514/c							
DEFINITION Human papillomavirus clone HPV1627 E6 protein (E6) gene, partial							
cds.							
U14514							
U14514.1 GI:984949							
Accession							
VERSION							
KEYWORDS							
SOURCE							
ORGANISM							
Human papillomavirus							
Human papillomavirus							
viruses; dsDNA viruses, no RNA stage; Papillomaviridae;							
unclassified Papillomaviridae.							
1 (bases 1 to 273)							
Haegert,D.G., Galutira,D.F. and Youngusband,B.H.							
Sequence variation in the E6 gene of human papillomavirus type 16							
Unpublished							
2 (bases 1 to 273)							
Galutira,D.F.							
Direct Submission							
Submitted (07-SEP-1994) Dante F. Galutira, Memorial University of							
Newfoundland, Discipline of Pathology, Fac. Medicine, St. John's,							
Newfoundland, A1B 3V6, Canada							
Location/Qualifiers							
1..273							
/organism="Human papillomavirus"							
/mol_type="genomic DNA"							
/strain="HPV16"							
/db_xref="taxon:10566"							
/clone="HPV1627"							
/tissue_type="cervical tissue from invasive carcinoma"							
1..273							
/gene="E6"							
<1..>273							
/gene="E6"							
/codon_start=2							
/product="E6 protein"							
/protein_id="AAB60568.2"							

/db_xref="GI:12831649"
/translation="VYDPAFRLDIYRDGNPYAVCDKCLKFYSKISEYRYCYSLYG
TTLEQQYNKPLCDLIRINCQKPLCPBEKQRLDKQRFHNRGR"
61
/gene="B6"
/note="T to A transversion from PA16, GenBank Accession
Number K02718, position 286; NCBI gi: 333031"
/replace="t"
64
/gene="B6"
/note="A to G transition from PA16, GenBank Accession
Number K02718, position 289; NCBI gi: 333031"
/replace="a"
110
/gene="B6"
/note="C to T transition from PA16, GenBank Accession
Number K02718, position 335; NCBI gi: 333031"
/replace="c"
variation

variation

ORIGIN

Query Match 100.0%; Score 21; DB 10; Length 273;
Best Local Similarity 100.0%; Pred. No. 0.3;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CAGGACACAGTGGCTTTTGAC 21
|||||
215 CAGGACACAGTGGCTTTTGAC 195

Db 215 CAGGACACAGTGGCTTTTGAC 195

RESULT 13
HPV14516 274 bp DNA linear VRL 14-SEP-1995
LOCUS Human papillomavirus clone HPV1649 E6 protein (E6) gene, partial
DEFINITION cds.
ACCESSION U14516
VERSION U14516.1 GI:984953
KEYWORDS Human papillomavirus
SOURCE Human papillomavirus
ORGANISM Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
unclassified Papillomaviridae.
REFERENCE 1 (bases 1 to 274)
Haeger, D.G., Galutira, D.F. and Younghusband, B.H.
TITLE Sequence variation in the E6 gene of human papillomavirus type 16
JOURNAL Unpublished
AUTHORS 2 (bases 1 to 274)
Galutira, D.F.
TITLES Direct Submission
JOURNAL Submitted (07-SEP-1994) Dante F. Galutira, Memorial University of
Newfoundland, Discipline of Pathology, Fac. Medicine, St. John's,
Newfoundland, A1B 3X6, Canada
FEATURES
source
1.274
/organism="Human papillomavirus"
/mol_type="genomic DNA"
/strain="HPV16"
/db_xref="taxon:10566"
/clone="HPV1649"
/issue_type="cervical tissue from invasive carcinoma"
1.274
/gene="B6"
/gene="B6"
/gene="B6"
/gene="B6"
/codon_start=2
/product="E6 protein"
/protein_id="AAB60570.1"
/db_xref="GI:984954"
/translation="VYDPAFRLDIYRDGNPYAVCDKCLKFYSKISEYRYCYSLYG
TTLEQQYNKPLCDLIRINCQKPLCPBEKQRLDKQRFHNRGR"
61
/gene="B6"
/note="T to A transversion from PA16, GenBank Accession
Number K02718, position 286; NCBI gi: 333031"

/replace="t"
64
/gene="B6"
/note="A to G transition from PA16, GenBank Accession
Number K02718, position 289; NCBI gi: 333031"
/replace="a"
110
/gene="B6"
/note="C to T transition from PA16, GenBank Accession
Number K02718, position 335; NCBI gi: 333031"
/replace="c"
125
/gene="B6"
/note="T to G transversion from PA16, GenBank Accession
Number K02718, position 350; NCBI gi: 333031"
/replace="t"
variation

variation

ORIGIN

Query Match 100.0%; Score 21; DB 10; Length 274;
Best Local Similarity 100.0%; Pred. No. 0.3;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CAGGACACAGTGGCTTTTGAC 21
|||||
215 CAGGACACAGTGGCTTTTGAC 195

Db 215 CAGGACACAGTGGCTTTTGAC 195

RESULT 14
AP404704 298 bp DNA linear VRL 09-OCT-2003
LOCUS Human papillomavirus type 16 isolate HPV166C13 E6 protein (E6)
DEFINITION gene, partial cds.
ACCESSION AP404704
VERSION AP404704.1 GI:15529605
KEYWORDS Human papillomavirus type 16
SOURCE Human papillomavirus type 16
ORGANISM Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
Alphapapillomavirus.
REFERENCE 1 (bases 1 to 298)
Wates, K.J., Thompson, C.H., Coscart, Y.E. and Rose, B.R.
TITLE Sequence variation and physical state of human papillomavirus type
16 cervical cancer isolates from Australia and New Caledonia
JOURNAL Int. J. Cancer 97 (6), 868-874 (2002)
PUBMED 11857370
REFERENCE 2 (bases 1 to 298)
Wates, K.J., Thompson, C.H., Coscart, Y.E. and Rose, B.R.
TITLES Direct Submission
JOURNAL Submitted (31-OCT-2001) Department of Infectious Diseases,
University of Sydney, Blackburn Building, D06, Off Western Avenue,
Camperdown, Sydney, New South Wales 2006, Australia
FEATURES
source
1.298
/organism="Human papillomavirus type 16"
/mol_type="genomic DNA"
/isolate="HPV166C13"
/db_xref="taxon:333760"
/gene="B6"
/gene="B6"
/gene="B6"
/gene="B6"
/codon_start=2
/product="E6 protein"
/protein_id="AA101365.1"
/db_xref="GI:15529606"
/translation="DITLFCVCKQQLLRBYVDPAFRLDIYRDGNPYAVCDKCL
FYSKISEYRYCYSLYGTLEQQYNKPLCDLIRINCQKPLCPBEKQRLDKQ"
61
ORIGIN

Query Match 100.0%; Score 21; DB 10; Length 298;
Best Local Similarity 100.0%; Pred. No. 0.3;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CAGGACACAGTGGCTTTGAC 21
|||||
Db 266 CAGGACACAGTGGCTTTGAC 246

RESULT 15
AF404692/c 370 bp DNA linear VRL 09-OCT-2003
LOCUS
DEFINITION Human papillomavirus type 16 isolate HPV16B6CC1 E6 protein (E6) and
E7 protein (E7) genes, partial cds.
ACCESSION AF404692
VERSION AF404692.1 GI:15529569
KEYWORDS
SOURCE Human papillomavirus type 16
ORGANISM Human papillomavirus type 16
VIRUSES; dsDNA viruses, no RNA stage; Papillomaviridae;
Alphapapillomavirus.
REFERENCE
AUTHORS 1 (bases 1 to 370)
Watts,K.J., Thompson,C.H., Cossart,Y.E. and Rose,B.R.
TITLE Sequence variation and physical state of human papillomavirus type
16 cervical cancer isolates from Australia and New Caledonia
JOURNAL Int. J. Cancer 97 (6), 868-874 (2002)
PUBMED 11857370
REFERENCES 2 (bases 1 to 370)
Watts,K.J., Thompson,C.H., Cossart,Y.E. and Rose,B.R.
AUTHORS Direct Submission
JOURNAL Submitted (31-JUL-2001) Department of Infectious Diseases,
University of Sydney, Blackburn Building, D06, Off Western Avenue,
Camperdown, Sydney, New South Wales 2006, Australia
Location/Qualifiers
FEATURES
source 1..370
/organism="Human papillomavirus type 16"
/mol_type="genomic DNA"
/isolate="HPV16B6CC1"
/db_xref="taxon:333760"
<1..313
/gene="E6"
<1..313
/gene="E6"
/codon_start=2
/product="E6 protein"
/protein_id="AAL01342.1"
/db_xref="GI:15529571"
/translation="DLICIVRDGDPVAVCDKCKEYSKISRYHYCVSLYGTLEQF
NKPICDLIRICINCQKPLCEBEKQRHLDKQKFRINIGRWTCMSCCRSRRTREIQ
L"
316..>370
/gene="E7"
316..>370
/gene="E7"
/codon_start=1
/product="E7 protein"
/protein_id="AAL01341.1"
/db_xref="GI:15529570"
/translation="MHGDTPTLHRYMLDQPE"

ORIGIN
Query Match 100.0%; Score 21; DB 10; Length 370;
Best Local Similarity 100.0%; Pred. No. 0.3;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CAGGACACAGTGGCTTTGAC 21
|||||
Db 194 CAGGACACAGTGGCTTTGAC 174

RESULT 16
S51110/c 421 bp DNA linear VRL 08-MAY-1993
LOCUS
DEFINITION off B6 (human papillomavirus HPV, type 16, head and neck tumor,
S51110
Genomic, 421 nt).
ACCESSION S51110
VERSION S51110.1 GI:262061

KEYWORDS
SOURCE Human papillomavirus
ORGANISM Human papillomavirus
VIRUSES; dsDNA viruses, no RNA stage; Papillomaviridae;
unclassified Papillomaviridae.
REFERENCE
AUTHORS 1 (bases 1 to 421)
Tyan,Y.S., Liu,S.T., Ong,W.R., Chen,M.L., Shu,C.H. and Chang,Y.S.
TITLE Detection of Epstein-Barr virus and human papillomavirus in head
and neck tumors
JOURNAL J. Clin. Microbiol. 31 (1), 53-56 (1993)
PUBMED 8380183
REFERENCES GenBank staff at the National Library of Medicine created this
entry [NCBI gisdbg 120826] from the original journal article.
COMMENT PCR primer sequences: 1-21 and 400-421.
FEATURES
source 1..421
/organism="Human papillomavirus"
/mol_type="genomic DNA"
/db_xref="taxon:10566"
/note="type: 16"
1..421
/gene="orf E6"

ORIGIN
Query Match 100.0%; Score 21; DB 10; Length 421;
Best Local Similarity 100.0%; Pred. No. 0.3;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CAGGACACAGTGGCTTTGAC 21
|||||
Db 338 CAGGACACAGTGGCTTTGAC 318

RESULT 17
AF404695/c 451 bp DNA linear VRL 09-OCT-2003
LOCUS
DEFINITION Human papillomavirus type 16 isolate HPV16B6CC4 E6 protein (E6) and
E7 protein (E7) genes, partial cds.
ACCESSION AF404695
VERSION AF404695.1 GI:15529578
KEYWORDS
SOURCE Human papillomavirus type 16
ORGANISM Human papillomavirus type 16
VIRUSES; dsDNA viruses, no RNA stage; Papillomaviridae;
Alphapapillomavirus.
REFERENCE
AUTHORS 1 (bases 1 to 451)
Watts,K.J., Thompson,C.H., Cossart,Y.E. and Rose,B.R.
TITLE Sequence variation and physical state of human papillomavirus type
16 cervical cancer isolates from Australia and New Caledonia
JOURNAL Int. J. Cancer 97 (6), 868-874 (2002)
PUBMED 11857370
REFERENCES 2 (bases 1 to 451)
Watts,K.J., Thompson,C.H., Cossart,Y.E. and Rose,B.R.
AUTHORS Direct Submission
JOURNAL Submitted (31-JUL-2001) Department of Infectious Diseases,
University of Sydney, Blackburn Building, D06, Off Western Avenue,
Camperdown, Sydney, New South Wales 2006, Australia
Location/Qualifiers
FEATURES
source 1..451
/organism="Human papillomavirus type 16"
/mol_type="genomic DNA"
/isolate="HPV16B6CC4"
/db_xref="taxon:333760"
<1..395
/gene="E6"
<1..395
/gene="E6"
/codon_start=3
/product="E6 protein"
/protein_id="AAL01347.1"
/db_xref="GI:15529579"
/translation="TTHDILTECVCKQQLRREYVGFAPRDLCTIVRDGPNPAVCDK
CLKFYSKISRYHYCVSLYGTLEQYNKPLCDLIRICINCQKPLCEBEKQRHLDKQKQ

gene RPHNIRGWTGRCMSCRSSRTRRETOL"
398..>451
/gene="E7"
398..>451
CDS /gene="E7"
/codon_start=1
/product="E7 protein"
/protein_id="AL01348.1"
/db_xref="GI:15529580"
/translation="MHGDTPTLHEVMDLQPS"

ORIGIN

Query Match 100.0%; Score 21; DB 10; Length 451;
Best Local Similarity 100.0%; Pred. No. 0.3;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAGGACACAGTGGCTTTTGAC 21
|||||
Db 276 CAGGACACAGTGGCTTTTGAC 256

RESULT 18
AP404696/c 451 bp DNA linear VRL 09-OCT-2003
LOCUS Human papillomavirus type 16 isolate HPV16B6CC5 E6 protein (B6) and
DEFINITION E7 protein (E7) genes, partial cds.
ACCESSION AF404696
VERSION AF404696.1 GI:15529581
KEYWORDS Human papillomavirus type 16
SOURCE Human papillomavirus type 16
ORGANISM Viruses; daDNA viruses, no RNA stage; Papillomaviridae;
Alphapapillomavirus.
1 (bases 1 to 451)

REFERENCE
AUTHORS Watts,K.J., Thompson,C.H., Coscart,Y.E. and Rose,B.R.
TITLE Sequence variation and physical state of human papillomavirus type
JOURNAL 16 cervical cancer isolates from Australia and New Caledonia
PUBMED 11857370
AUTHORS Int. J. Cancer 97 (6), 868-874 (2002)
AUTHORS 2 (bases 1 to 451)
TITLE Watts,K.J., Thompson,C.H., Coscart,Y.E. and Rose,B.R.
JOURNAL Direct Submission
PUBMED Submitted (31-JUL-2001) Department of Infectious Diseases,
AUTHORS University of Sydney, Blackburn Building, D06, Off Western Avenue,
JOURNAL Camperdown, Sydney, New South Wales 2006, Australia

FEATURES
source Location/Qualifiers
1..451
/organism="Human papillomavirus type 16"
/mol_type="genomic DNA"
/isolate="HPV16B6CC5"
/db_xref="taxon:333760"
1..395
/gene="E6"
1..395
/gene="E6"
/codon_start=3
/product="E6 protein"
/protein_id="AA01349.1"
/db_xref="GI:15529582"
/translation="TTHNIIILCCVYKQQLRREVDPAFRDICTYRRGNPYAVCDK
CKFKYSKISBYRHYCYSVYGTLEQQYNPLCDLIRTCINQKPLCPBESKORLDDKKQ
RPHNIRGWTGRCMSCRSSRTRRETOL"

gene
CDS

gene RPHNIRGWTGRCMSCRSSRTRRETOL"
398..>451
/gene="E7"
398..>451
CDS /gene="E7"
/codon_start=1
/product="E7 protein"
/protein_id="AA01350.1"
/db_xref="GI:15529583"
/translation="MHGDTPTLHEVMDLQPS"

ORIGIN

Query Match 100.0%; Score 21; DB 10; Length 451;
Best Local Similarity 100.0%; Pred. No. 0.3;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAGGACACAGTGGCTTTTGAC 21
|||||
Db 276 CAGGACACAGTGGCTTTTGAC 256

RESULT 19
AP404699/c 451 bp DNA linear VRL 09-OCT-2003
LOCUS Human papillomavirus type 16 isolate HPV16B6CC8 E6 protein (B6) and
DEFINITION E7 protein (E7) genes, partial cds.
ACCESSION AF404699
VERSION AF404699.1 GI:15529590
KEYWORDS Human papillomavirus type 16
SOURCE Human papillomavirus type 16
ORGANISM Viruses; daDNA viruses, no RNA stage; Papillomaviridae;
Alphapapillomavirus.
1 (bases 1 to 451)

REFERENCE
AUTHORS Watts,K.J., Thompson,C.H., Coscart,Y.E. and Rose,B.R.
TITLE Sequence variation and physical state of human papillomavirus type
JOURNAL 16 cervical cancer isolates from Australia and New Caledonia
PUBMED 11857370
AUTHORS Int. J. Cancer 97 (6), 868-874 (2002)
AUTHORS 2 (bases 1 to 451)
TITLE Watts,K.J., Thompson,C.H., Coscart,Y.E. and Rose,B.R.
JOURNAL Direct Submission
PUBMED Submitted (31-JUL-2001) Department of Infectious Diseases,
AUTHORS University of Sydney, Blackburn Building, D06, Off Western Avenue,
JOURNAL Camperdown, Sydney, New South Wales 2006, Australia

FEATURES
source Location/Qualifiers
1..451
/organism="Human papillomavirus type 16"
/mol_type="genomic DNA"
/isolate="HPV16B6CC8"
/db_xref="taxon:333760"
1..395
/gene="E6"
1..395
/gene="E6"
/codon_start=3
/product="E6 protein"
/protein_id="AA01355.1"
/db_xref="GI:15529591"
/translation="TTHNIIILCCVYKQQLRREVDPAFRDICTYRRGNPYAVCDK
CKFKYSKISBYRHYCYSVYGTLEQQYNPLCDLIRTCINQKPLCPBESKORLDDKKQ
RPHNIRGWTGRCMSCRSSRTRRETOL"

gene
CDS

gene RPHNIRGWTGRCMSCRSSRTRRETOL"
398..>451
/gene="E7"
398..>451
CDS /gene="E7"
/codon_start=1
/product="E7 protein"
/protein_id="AA01356.1"
/db_xref="GI:15529592"
/translation="MHGDTPTLHEVMDLHPS"

ORIGIN

Query Match 100.0%; Score 21; DB 10; Length 451;
Best Local Similarity 100.0%; Pred. No. 0.3;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAGGACACAGTGGCTTTTGAC 21
|||||
Db 276 CAGGACACAGTGGCTTTTGAC 256

RESULT 20
AP404700/c 451 bp DNA linear VRL 09-OCT-2003
LOCUS

DEFINITION Human papillomavirus type 16 isolate HPV16B6CC9 E6 protein (E6) and
AF404700
AF404700.1 GI:15529593
VERSION
KEYWORDS Human papillomavirus type 16
SOURCE Human papillomavirus type 16
ORGANISM Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
Alphapapillomavirus.
REFERENCE
AUTHORS 1 (bases 1 to 451)
TITLE Waites,K.J., Thompson,C.H., Cossart,Y.E. and Rose,B.R.
JOURNAL Sequence variation and physical state of human papillomavirus type
PUBMED 16 cervical cancer isolates from Australia and New Caledonia
11857370
2 (bases 1 to 451)
REFERENCE Waites,K.J., Thompson,C.H., Cossart,Y.E. and Rose,B.R.
AUTHORS Direct Submission
JOURNAL Submitted (31-JUL-2001) Department of Infectious Diseases,
University of Sydney, Blackburn Building, D06, Off Western Avenue,
Camperdown, Sydney, New South Wales 2006, Australia
Location/Qualifiers
1. .451
/organism="Human papillomavirus type 16"
/mol_type="genomic DNA"
/isolate="HPV16B6CC9"
/db_xref="taxon:333760"
<1. .395
/gene="E6"
<1. .395
/gene="E6"
/codon_start=3
/product="E6 protein"
/protein_id="AA01357.1"
/db_xref="GI:15529594"
/translation="TIDHILIECYCKQQLRREVDPAFADLCIVYRDGNPYAVCDK
CLKPKYSISRHYCYGVGTTLLEQYNKPCDLLIRINCQKPLCEPEKQRHDKQ
RFNIRGRMTGRMCCRSRSTRRETQL"
398. .>451
/gene="E7"
398. .>451
/gene="E7"
/codon_start=1
/product="E7 protein"
/protein_id="AA01358.1"
/db_xref="GI:15529595"
/translation="MHGDTPTLHRYMLDQPE"

ORIGIN
Query Match 100.0%; Score 21; DB 10; Length 451;
Best Local Similarity 100.0%; Pred. No. 0.3;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 CAGGACACAGTGGCTTTGAC 21
|||
276 CAGGACACAGTGGCTTTGAC 256

Db

RESULT 21
AF404701/c 451 bp DNA linear VRL 09-OCT-2003
LOCUS
DEFINITION Human papillomavirus type 16 isolate HPV16B6CC10 E6 protein (E6)
and E7 protein (E7) genes, partial cds.
ACCESSION AF404701
AF404701.1 GI:15529596
VERSION
KEYWORDS Human papillomavirus type 16
SOURCE Human papillomavirus type 16
ORGANISM Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
Alphapapillomavirus.
REFERENCE
AUTHORS 1 (bases 1 to 451)
TITLE Waites,K.J., Thompson,C.H., Cossart,Y.E. and Rose,B.R.
JOURNAL Sequence variation and physical state of human papillomavirus type

JOURNAL 16 cervical cancer isolates from Australia and New Caledonia
PUBMED Int. J. Cancer 97 (6), 868-874 (2002)
11857370
2 (bases 1 to 451)
REFERENCE Waites,K.J., Thompson,C.H., Cossart,Y.E. and Rose,B.R.
AUTHORS Direct Submission
JOURNAL Submitted (31-JUL-2001) Department of Infectious Diseases,
University of Sydney, Blackburn Building, D06, Off Western Avenue,
Camperdown, Sydney, New South Wales 2006, Australia
Location/Qualifiers
1. .451
/organism="Human papillomavirus type 16"
/mol_type="genomic DNA"
/isolate="HPV16B6CC10"
/db_xref="taxon:333760"
<1. .395
/gene="E6"
<1. .395
/gene="E6"
/codon_start=3
/product="E6 protein"
/protein_id="AA01359.1"
/db_xref="GI:15529597"
/translation="PIHDILIECYCKQQLRREVDPAFADLCIVYRDGNPYAVCDK
CLKPKYSISRHYCYGVGTTLLEQYNKPCDLLIRINCQKPLCEPEKQRHDKQ
RFNIRGRMTGRMCCRSRSTRRETQL"
398. .>451
/gene="E7"
398. .>451
/gene="E7"
/codon_start=1
/product="E7 protein"
/protein_id="AA01360.1"
/db_xref="GI:15529598"
/translation="MHGDTPTLHRYMLDQPE"

ORIGIN
Query Match 100.0%; Score 21; DB 10; Length 451;
Best Local Similarity 100.0%; Pred. No. 0.3;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 CAGGACACAGTGGCTTTGAC 21
|||
276 CAGGACACAGTGGCTTTGAC 256

Db

RESULT 22
AF404703/c 451 bp DNA linear VRL 09-OCT-2003
LOCUS
DEFINITION Human papillomavirus type 16 isolate HPV16B6CC12 E6 protein (E6)
and E7 protein (E7) genes, partial cds.
ACCESSION AF404703
AF404703.1 GI:15529602
VERSION
KEYWORDS Human papillomavirus type 16
SOURCE Human papillomavirus type 16
ORGANISM Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
Alphapapillomavirus.
REFERENCE
AUTHORS 1 (bases 1 to 451)
TITLE Waites,K.J., Thompson,C.H., Cossart,Y.E. and Rose,B.R.
JOURNAL Sequence variation and physical state of human papillomavirus type
PUBMED 16 cervical cancer isolates from Australia and New Caledonia
11857370
2 (bases 1 to 451)
REFERENCE Waites,K.J., Thompson,C.H., Cossart,Y.E. and Rose,B.R.
AUTHORS Direct Submission
JOURNAL Submitted (31-JUL-2001) Department of Infectious Diseases,
University of Sydney, Blackburn Building, D06, Off Western Avenue,
Camperdown, Sydney, New South Wales 2006, Australia
Location/Qualifiers
1. .451
/organism="Human papillomavirus type 16"

gene /mol_type="genomic DNA"
/isolate="HPV16B6C12"
/db_xref="taxon:333760"
1..395
/gene="E6"
1..395
/gene="E6"
/codon_start=3
/product="E6 protein"
/protein_id="AA01363.1"
/db_xref="GI:15529603"
/translation="TTHNITLCEVCYCKQQLLRVYDPAFPRDICTIYRGNPAVCDK
CINFTSKISEYRHYCYSLYGTLSQQYNKPLCDLIRICINCKPILCEBKRRLDKKQ
RFHNIIRGWTGRCMSCCRSSRRTRRETOL"
398..>451
/gene="E7"
398..>451
/gene="E7"
/codon_start=1
/product="E7 protein"
/protein_id="AA01364.1"
/db_xref="GI:15529604"
/translation="MHGDTPTLHEVWLIDLOPE"
/translaction="MHGDTPTLHEVWLIDLOPE"
ORIGIN

Query Match 100.0%; Score 21; DB 10; Length 451;
Best Local Similarity 100.0%; Pred. No. 0.3;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAGGACACAGTGGCTTTTGAC 21
|||||
Db 276 CAGGACACAGTGGCTTTTGAC 256

RESULT 23
AR167393/c
LOCUS AR167393 456 bp DNA linear PAT 17-DEC-2001
DEFINITION Sequence 29 from patent US 6287569.
ACCESSION AR167393
VERSION AR167393.1 GI:17903171
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 456)
AUTHORS Kipps, T.J. and Wu, Y.
TITLE Vaccines with enhanced intracellular processing
JOURNAL Patent: US 6287569-A 29 11-SEP-2001;
FEATURES
source Location/Qualifiers
1..456
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN

Query Match 100.0%; Score 21; DB 2; Length 456;
Best Local Similarity 100.0%; Pred. No. 0.3;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAGGACACAGTGGCTTTTGAC 21
|||||
Db 337 CAGGACACAGTGGCTTTTGAC 317

RESULT 24
AR177943/c
LOCUS AR177943 456 bp DNA linear PAT 17-DEC-2001
DEFINITION Sequence 6 from patent US 6313373.
ACCESSION AR177943
VERSION AR177943.1 GI:17920298
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
FEATURES
source Location/Qualifiers
1..456
/organism="unknown"
/mol_type="unassigned DNA"

REFERENCE 1 (bases 1 to 456)
AUTHORS Eckert, R.L. and Crish, J.F.
TITLE Tissue specific promoters and transgenic mouse for the screening of
pharmaceuticals
JOURNAL Patent: US 6313373-A 6 06-NOV-2001;
FEATURES
source Location/Qualifiers
1..456
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN

Query Match 100.0%; Score 21; DB 2; Length 456;
Best Local Similarity 100.0%; Pred. No. 0.3;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAGGACACAGTGGCTTTTGAC 21
|||||
Db 337 CAGGACACAGTGGCTTTTGAC 317

RESULT 25
AF327851/c
LOCUS AF327851 456 bp DNA linear VRL 04-JAN-2001
DEFINITION Human papillomavirus type 16 early transforming protein E6 variant
ACCESSION AF327851
VERSION AF327851.1 GI:12025467
KEYWORDS
SOURCE Human papillomavirus type 16
ORGANISM Human papillomavirus type 16
REFERENCE 1 (bases 1 to 456)
AUTHORS Zhang, F., Qian, D., Ma, J., Lin, R., Ming, W., Zhong, Z., Zhang, Q. and
Cloning, F. and Sequencing of HPV16 E6 gene from Cervical Carcinoma
Biopsies in Xinjiang
JOURNAL Shengwu Huaxue Yu Shengwu Wuli Jinzhan (2001) In press
ACCESSION 2 (bases 1 to 456)
AUTHORS Ma, Z. and Zhang, F.
TITLE Direct Submision
JOURNAL Submitted (12-DEC-2000) Department of Biology, University of
Xinjiang, 14 Shengqi Road, Urumqi, Xinjiang 830046, P.R. China
FEATURES
source Location/Qualifiers
1..456
/organism="Human papillomavirus type 16"
/proviral
/mol_type="genomic DNA"
/strain="Xinjiang"
/specific host="Homo sapiens"
/db_xref="taxon:333760"
/note="HPV16"
1..456
/gene="E6"
1..456
/gene="E6"
/codon_start=1
/product="early transforming protein E6 variant"
/protein_id="AAG45940.1"
/db_xref="GI:12025468"
/translation="MFQDQERPRKLPQCLTSLQTTIHILLCEVCYCKQQLLRVYD
FAFRDICTIYRGNPAVCDKCLKFTSKISEYRHYCYSLYGTLSQQYNKPLCDLIRICINCKPILCEBKRRLDKKORFHNIRGWTGRCMSCCRSSRRTRRETOL"

ORIGIN

Query Match 100.0%; Score 21; DB 10; Length 456;
Best Local Similarity 100.0%; Pred. No. 0.3;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAGGACACAGTGGCTTTTGAC 21
|||||
Db 337 CAGGACACAGTGGCTTTTGAC 317

RESULT 26
HPU34107/c 456 bp DNA linear VRL 08-MAR-1996
LOCUS Human papillomavirus type 16, isolate NM 4094, early transforming
DEFINITION protein E6 (E6) gene, complete cds.
ACCESSION U34107
VERSION U34107.1 GI:10987219
KEYWORDS
SOURCE Human papillomavirus
ORGANISM Human papillomavirus
Virusess; dsDNA viruses, no RNA stage; Papillomaviridae;
unclassified Papillomaviridae.
REFERENCE
AUTHORS Yamada,T., Wheeler,C.M., Halpern,A.L., Stewart,A.C., Hildesheim,A.
and Jenison,S.A.
TITLE Human papillomavirus type 16 variant lineages in United States
JOURNAL populations characterized by nucleotide sequence analysis of the
PUBMED E6, I2, and I1 coding segments
7494284 J. Virol. 69 (12), 7743-7753 (1995)
2 (bases 1 to 456)
REFERENCE Farmer,A.D.
AUTHORS Direct Submission
TITLE Submitted (16-AUG-1995) Andrew D. Farmer, Los Alamos National
JOURNAL Laboratory, HIV Sequence Database, T-10, Mail Stop K710, Los
Alamos, NM 87501, USA
FEATURES
source Location/Qualifiers
1. .456
/organism="Human papillomavirus"
/mol_type="genomic DNA"
/isolate="NM 4094"
/db_xref="taxon:10566"
/note="collected by cervical swab at the University of New
Mexico, New Mexico, United States."
gene 1. .456
/gene="E6"
1. .456
/gene="E6"
/codon_start=1
/product="early transforming protein E6"
/protein_id="AA91654.1"
/db_xref="GI:1098720"
/translation="MFQDPOBRPKLPHLCTELQTHDITLIEGVCKQQLRREYD
FARPDCTIVRDGNPVAVCDCLEKFSKISRYCYSGVGTTLSEQYNKPLCDLLR
CINQCKPLCEBKORHLDKORFNHNGRWTCMSCSCSSRTRETOL"
ORIGIN
Query Match 100.0%; Score 21; DB 10; Length 456;
Best Local Similarity 100.0%; Pred. No. 0.3;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CAGGACACAGTGGCTTTGAC 21
|||||
Db 337 CAGGACACAGTGGCTTTGAC 317
RESULT 27
HPU34108/c 456 bp DNA linear VRL 08-MAR-1996
LOCUS Human papillomavirus type 16, isolate NM 9999, early transforming
DEFINITION protein E6 (E6) gene, complete cds.
ACCESSION U34108
VERSION U34108.1 GI:1098721
KEYWORDS
SOURCE Human papillomavirus
ORGANISM Human papillomavirus
Virusess; dsDNA viruses, no RNA stage; Papillomaviridae;
unclassified Papillomaviridae.
REFERENCE
AUTHORS Yamada,T., Wheeler,C.M., Halpern,A.L., Stewart,A.C., Hildesheim,A.
and Jenison,S.A.
TITLE Human papillomavirus type 16 variant lineages in United States

JOURNAL populations characterized by nucleotide sequence analysis of the
PUBMED E6, I2, and I1 coding segments
7494284 J. Virol. 69 (12), 7743-7753 (1995)
2 (bases 1 to 456)
REFERENCE Farmer,A.D.
AUTHORS Direct Submission
TITLE Submitted (16-AUG-1995) Andrew D. Farmer, Los Alamos National
JOURNAL Laboratory, HIV Sequence Database, T-10, Mail Stop K710, Los
Alamos, NM 87501, USA
FEATURES
source Location/Qualifiers
1. .456
/organism="Human papillomavirus"
/mol_type="genomic DNA"
/isolate="NM 9999"
/db_xref="taxon:10566"
/note="collected by cervical swab at the University of New
Mexico, New Mexico, United States."
gene 1. .456
/gene="E6"
1. .456
/gene="E6"
/codon_start=1
/product="early transforming protein E6"
/protein_id="AA91655.1"
/db_xref="GI:1098723"
/translation="MFQDPOBRPKLPHLCTELQTHDITLIEGVCKQQLRREYD
FARPDCTIVRDGNPVAVCDCLEKFSKISRYCYSGVGTTLSEQYNKPLCDLLR
CINQCKPLCEBKORHLDKORFNHNGRWTCMSCSCSSRTRETOL"
ORIGIN
Query Match 100.0%; Score 21; DB 10; Length 456;
Best Local Similarity 100.0%; Pred. No. 0.3;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CAGGACACAGTGGCTTTGAC 21
|||||
Db 337 CAGGACACAGTGGCTTTGAC 317
RESULT 28
HPU34109/c 456 bp DNA linear VRL 08-MAR-1996
LOCUS Human papillomavirus type 16, isolate NM T197, early transforming
DEFINITION protein E6 (E6) gene, complete cds.
ACCESSION U34109
VERSION U34109.1 GI:1098723
KEYWORDS
SOURCE Human papillomavirus
ORGANISM Human papillomavirus
Virusess; dsDNA viruses, no RNA stage; Papillomaviridae;
unclassified Papillomaviridae.
REFERENCE
AUTHORS Yamada,T., Wheeler,C.M., Halpern,A.L., Stewart,A.C., Hildesheim,A.
and Jenison,S.A.
TITLE Human papillomavirus type 16 variant lineages in United States
JOURNAL populations characterized by nucleotide sequence analysis of the
PUBMED E6, I2, and I1 coding segments
7494284 J. Virol. 69 (12), 7743-7753 (1995)
2 (bases 1 to 456)
REFERENCE Farmer,A.D.
AUTHORS Direct Submission
TITLE Submitted (16-AUG-1995) Andrew D. Farmer, Los Alamos National
JOURNAL Laboratory, HIV Sequence Database, T-10, Mail Stop K710, Los
Alamos, NM 87501, USA
FEATURES
source Location/Qualifiers
1. .456
/organism="Human papillomavirus"
/mol_type="genomic DNA"
/isolate="NM T197"
/db_xref="taxon:10566"
/note="collected by cervical swab at the University of New

gene Mexico, New Mexico, United States."
1. .456
/gene="B6"
1. .456
/gene="B6"
/codon_start=1
/product="early transforming protein E6"
/protein_id="AA91656.1"
/db_xref="GI:1098724"
/translation="MFODPQERPRKLPOLCTELQTTIHIIIECYCKQQLRREYD
FARDICIVYRDGNPFAVCDKCLKFSKISRYHVCYSYGTLLDQYNNPLCDLLIR
CINCKRPLCEBKQRHLDKQRPHNIRGRWGTGCMSCRSRTRRETOL"

ORIGIN

Query Match 100.0%; Score 21; DB 10; Length 456;
Best Local Similarity 100.0%; Pred. No. 0.3;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAGGACACAGTGCGCTTTGAC 21
HPJ34110/c
LOCUS
DEFINITION Human papillomavirus type 16, isolate NM T446, early transforming
protein E6 (E6) gene, complete cds.
ACCESSION U34110.1
VERSION U34110.1 GI:1098725
KEYWORDS
SOURCE Human papillomavirus
ORGANISM Human papillomavirus
VIRUSES; dsDNA viruses, no RNA stage; Papillomaviridae;
unclassified Papillomaviridae.
1 (bases 1 to 456)
Yamada,T., Wheeler,C.M., Halpern,A.L., Stewart,A.C., Hildesheim,A.
and Jenison,S.A.
Human papillomavirus type 16 variant lineages in United States
populations characterized by nucleotide sequence analysis of the
E6, E2, and E1 coding segments
J. Virol. 69 (12), 7743-7753 (1995)
7494284
2 (bases 1 to 456)
Farmer,A.D.
Direct Submision
Submitted (16-AUG-1995) Andrew D. Farmer, Los Alamos National
Laboratory, HIV Sequence Database, T-10, Mail Stop K710, Los
Alamos, NM 87501, USA
Location/Qualifiers
1. .456
/organism="Human papillomavirus"
/mol_type="genomic DNA"
/isolate="NM T446"
/db_xref="taxon:10566"
/note="collected by cervical swab at the University of New
Mexico, New Mexico, United States."
1. .456
/gene="B6"
1. .456
/gene="B6"
/codon_start=1
/product="early transforming protein E6"
/protein_id="AA91657.1"
/db_xref="GI:1098726"
/translation="MFODPQERPRKLPOLCTELQTTIHIIIECYCKQQLRREYD
FARDICIVYRDGNPFAVCDKCLKFSKISRYHVCYSYGTLLDQYNNPLCDLLIR
CINCKRPLCEBKQRHLDKQRPHNIRGRWGTGCMSCRSRTRRETOL"

ORIGIN

Query Match 100.0%; Score 21; DB 10; Length 456;
Best Local Similarity 100.0%; Pred. No. 0.3;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAGGACACAGTGCGCTTTGAC 21
HPJ34110/c
LOCUS
DEFINITION Human papillomavirus type 16, isolate NM T455, early transforming
protein E6 (E6) gene, complete cds.
ACCESSION U34111.1
VERSION U34111.1 GI:1098727
KEYWORDS
SOURCE Human papillomavirus
ORGANISM Human papillomavirus
VIRUSES; dsDNA viruses, no RNA stage; Papillomaviridae;
unclassified Papillomaviridae.
1 (bases 1 to 456)
Yamada,T., Wheeler,C.M., Halpern,A.L., Stewart,A.C., Hildesheim,A.
and Jenison,S.A.
Human papillomavirus type 16 variant lineages in United States
populations characterized by nucleotide sequence analysis of the
E6, E2, and E1 coding segments
J. Virol. 69 (12), 7743-7753 (1995)
7494284
2 (bases 1 to 456)
Farmer,A.D.
Direct Submision
Submitted (16-AUG-1995) Andrew D. Farmer, Los Alamos National
Laboratory, HIV Sequence Database, T-10, Mail Stop K710, Los
Alamos, NM 87501, USA
Location/Qualifiers
1. .456
/organism="Human papillomavirus"
/mol_type="genomic DNA"
/isolate="NM T455"
/db_xref="taxon:10566"
/note="collected by cervical swab at the University of New
Mexico, New Mexico, United States."
1. .456
/gene="B6"
1. .456
/gene="B6"
/codon_start=1
/product="early transforming protein E6"
/protein_id="AA91658.1"
/db_xref="GI:1098728"
/translation="MFODPQERPRKLPOLCTELQTTIHIIIECYCKQQLRREYD
FARDICIVYRDGNPFAVCDKCLKFSKISRYHVCYSYGTLLDQYNNPLCDLLIR
CINCKRPLCEBKQRHLDKQRPHNIRGRWGTGCMSCRSRTRRETOL"

ORIGIN

Query Match 100.0%; Score 21; DB 10; Length 456;
Best Local Similarity 100.0%; Pred. No. 0.3;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAGGACACAGTGCGCTTTGAC 21
HPJ34112/c
LOCUS
DEFINITION Human papillomavirus type 16, isolate NM T529, early transforming
protein E6 (E6) gene, complete cds.
ACCESSION U34112.1
VERSION U34112.1 GI:1098729
KEYWORDS
SOURCE Human papillomavirus
ORGANISM Human papillomavirus
VIRUSES; dsDNA viruses, no RNA stage; Papillomaviridae;

unclassified Papillomaviridae.
REFERENCE 1 (bases 1 to 456)
AUTHORS Yamada,T., Wheeler,C.M., Halpern,A.L., Stewart,A.C., Hildesheim,A.
and Jenison,S.A.
TITLE Human papillomavirus type 16 variant lineages in United States
populations characterized by nucleotide sequence analysis of the
E6, E2, and L1 coding segments
J. Virol. 69 (12), 7743-7753 (1995)
JOURNAL 7494284
PUBMED 2 (bases 1 to 456)
REFERENCE Farmer,A.D.
AUTHORS Direct Submission
TITLE Submitted (16-AUG-1995) Andrew D. Farmer, Los Alamos National
Laboratory, HIV Sequence Database, T-10, Mail Stop K710, Los
Alamos, NM 87501, USA
FEATURES Location/Qualifiers
source 1. 456
/organism="Human papillomavirus"
/mol_type="genomic DNA"
/isolate="NM T529"
/db_xref="taxon:10566"
/note="collected by cervical swab at the University of New
Mexico, New Mexico, United States."
gene 1. 456
/gene="E6"
CDS 1. 456
/gene="E6"
/codon_start=1
/product="early transforming protein E6"
/protein_id="AA91659.1"
/db_xref="GI:1098730"
/translation="MFQDPQRPRLPHLCTBLQTTIHDIILCEVCCKOOLLREYVD
FARDLCIVRDGNPYAVCDKCLKFYSKISEYHYCYSLYGTTLBQYNNPCLDLIR
CINCQKPLCEBKQRHDKKQRFHNIRGRTGRCMSCCRSSRTRETYL"

Query Match 100.0%; Score 21; DB 10; Length 456;
Best Local Similarity 100.0%; Pred. No. 0.3;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CAGGACACAGTGGCTTTGAC 21
|||||
Db 337 CAGGACACAGTGGCTTTGAC 317

RESULT 32
HPJ34113/c 456 bp DNA linear VRL 08-MAR-1996
LOCUS Human papillomavirus type 16, isolate OR 0198, early transforming
DEFINITION protein E6 (E6) gene, complete cds.
ACCESSION U34113
VERSION U34113.1 GI:1098731
KEYWORDS Human papillomavirus
SOURCE Human papillomavirus
ORGANISM Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
unclassified Papillomaviridae.
REFERENCE 1 (bases 1 to 456)
AUTHORS Yamada,T., Wheeler,C.M., Halpern,A.L., Stewart,A.C., Hildesheim,A.
and Jenison,S.A.
TITLE Human papillomavirus type 16 variant lineages in United States
populations characterized by nucleotide sequence analysis of the
E6, E2, and L1 coding segments
J. Virol. 69 (12), 7743-7753 (1995)
JOURNAL 7494284
PUBMED 2 (bases 1 to 456)
REFERENCE Farmer,A.D.
AUTHORS Direct Submission
TITLE Submitted (16-AUG-1995) Andrew D. Farmer, Los Alamos National
Laboratory, HIV Sequence Database, T-10, Mail Stop K710, Los
Alamos, NM 87501, USA
FEATURES Location/Qualifiers
source 1. 456

/organism="Human papillomavirus"
/mol_type="genomic DNA"
/isolate="OR 0198"
/db_xref="taxon:10566"
/note="collected by cervical lavage at Kaiser Permanente,
Portland, Oregon, United States."
gene 1. 456
/gene="E6"
CDS 1. 456
/gene="E6"
/codon_start=1
/product="early transforming protein E6"
/protein_id="AA91660.1"
/db_xref="GI:1098732"
/translation="MFQDPQRPRLPHLCTBLQTTIHDIILCEVCCKOOLLREYVD
FARDLCIVRDGNPYAVCDKCLKFYSKISEYHYCYSLYGTTLBQYNNPCLDLIR
CINCQKPLCEBKQRHDKKQRFHNIRGRTGRCMSCCRSSRTRETYL"

Query Match 100.0%; Score 21; DB 10; Length 456;
Best Local Similarity 100.0%; Pred. No. 0.3;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CAGGACACAGTGGCTTTGAC 21
|||||
Db 337 CAGGACACAGTGGCTTTGAC 317

RESULT 33
HPJ34114/c 456 bp DNA linear VRL 08-MAR-1996
LOCUS Human papillomavirus type 16, isolate OR 1905, early transforming
DEFINITION protein E6 (E6) gene, complete cds.
ACCESSION U34114
VERSION U34114.1 GI:1098733
KEYWORDS Human papillomavirus
SOURCE Human papillomavirus
ORGANISM Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
unclassified Papillomaviridae.
REFERENCE 1 (bases 1 to 456)
AUTHORS Yamada,T., Wheeler,C.M., Halpern,A.L., Stewart,A.C., Hildesheim,A.
and Jenison,S.A.
TITLE Human papillomavirus type 16 variant lineages in United States
populations characterized by nucleotide sequence analysis of the
E6, E2, and L1 coding segments
J. Virol. 69 (12), 7743-7753 (1995)
JOURNAL 7494284
PUBMED 2 (bases 1 to 456)
REFERENCE Farmer,A.D.
AUTHORS Direct Submission
TITLE Submitted (16-AUG-1995) Andrew D. Farmer, Los Alamos National
Laboratory, HIV Sequence Database, T-10, Mail Stop K710, Los
Alamos, NM 87501, USA
FEATURES Location/Qualifiers
source 1. 456
/organism="Human papillomavirus"
/mol_type="genomic DNA"
/isolate="OR 1905"
/db_xref="taxon:10566"
/note="collected by cervical lavage at Kaiser Permanente,
Portland, Oregon, United States."
gene 1. 456
/gene="E6"
CDS 1. 456
/gene="E6"
/codon_start=1
/product="early transforming protein E6"
/protein_id="AA91661.1"
/db_xref="GI:1098734"
/translation="MFQDPQRPRLPHLCTBLQTTIHDIILCEVCCKOOLLREYVD
FARDLCIVRDGNPYAVCDKCLKFYSKISEYHYCYSLYGTTLBQYNNPCLDLIR
CINCQKPLCEBKQRHDKKQRFHNIRGRTGRCMSCCRSSRTRETYL"

JOURNAL Submitted (16-AUG-1995) Andrew D. Farmer, Los Alamos National Laboratory, HIV Sequence Database, T-10, Mail Stop K710, Los Alamos, NM 87501, USA

FEATURES
source
1. .456
/organism="Human papillomavirus"
/mol_type="genomic DNA"
/isolate="OR 3473"
/db_xref="taxon:10566"
/note="collected by cervical lavage at Kaiser Permanente, Portland, Oregon, United States."
1. .456
/gene="E6"
1. .456
/gene="E6"
/codon_start=1
/product="early transforming protein E6"
/protein_id="AA91664.1"
/db_xref="GI:1098740"
/translation="MFQDPQRPRIKLPDLCTELQTTIHDIILCEVCCKOQLRREYVD
FARPDLCIVRDGNPFAVCDKCKAFYSKISEYRYCYSYGTLEQQYNRPCLDLIR
CINQKPLCEBKQRHDKKQRFNIRGWTGRCMSCRCSRTRRETQL"

ORIGIN
Query Match 100.0%; Score 21; DB 10; Length 456;
Best Local Similarity 100.0%; Pred. No. 0.3; Mismatches 0; Indels 0; Gaps 0;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAGGACACAGTGGCTTTGAC 21
|||||
337 CAGGACACAGTGGCTTTGAC 317

RESULT 37
HPV34118/c 456 bp DNA linear VRL 08-MAR-1996
LOCUS Human papillomavirus type 16, isolate OR 3759, early transforming
DEFINITION protein E6 (E6) gene, complete cds.
VERSION U34118.1 GI:1098741
KEYWORDS Human papillomavirus
SOURCE Human papillomavirus
ORGANISM Human papillomavirus
Virusess; dsDNA viruses, no RNA stage; Papillomaviridae;
unclassified Papillomaviridae.
1 (bases 1 to 456)
Yamada,T., Wheeler,C.M., Halpern,A.L., Stewart,A.C., Hildesheim,A.
and Jenison,S.A.
Human papillomavirus type 16 variant lineages in United States
populations characterized by nucleotide sequence analysis of the
E6, E2, and L1 coding segments
J. Virol. 69 (12), 7743-7753 (1995)
7494284
2 (bases 1 to 456)
Farmer,A.D.
Direct Submission
Submitted (16-AUG-1995) Andrew D. Farmer, Los Alamos National
Laboratory, HIV Sequence Database, T-10, Mail Stop K710, Los
Alamos, NM 87501, USA
location/Qualifiers
1. .456
/organism="Human papillomavirus"
/mol_type="genomic DNA"
/isolate="OR 3759"
/db_xref="taxon:10566"
/note="collected by cervical lavage at Kaiser Permanente,
Portland, Oregon, United States."
1. .456
/gene="E6"
1. .456
/gene="E6"
/codon_start=1
/product="early transforming protein E6"

gene
1. .456
/gene="E6"
1. .456
/gene="E6"
/codon_start=1
/product="early transforming protein E6"

CDS
1. .456
/gene="E6"
1. .456
/gene="E6"
/codon_start=1
/product="early transforming protein E6"

/protein_id="AA91665.1"
/db_xref="GI:1098742"
/translation="MFQDPQRPRIKLPDLCTELQTTIHDIILCEVCCKOQLRREYVD
FARPDLCIVRDGNPFAVCDKCKAFYSKISEYRYCYSYGTLEQQYNRPCLDLIR
CINQKPLCEBKQRHDKKQRFNIRGWTGRCMSCRCSRTRRETQL"

ORIGIN
Query Match 100.0%; Score 21; DB 10; Length 456;
Best Local Similarity 100.0%; Pred. No. 0.3; Mismatches 0; Indels 0; Gaps 0;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAGGACACAGTGGCTTTGAC 21
|||||
337 CAGGACACAGTGGCTTTGAC 317

RESULT 38
HPV34119/c 456 bp DNA linear VRL 08-MAR-1996
LOCUS Human papillomavirus type 16, isolate OR 4541, early transforming
DEFINITION protein E6 (E6) gene, complete cds.
VERSION U34119.1 GI:1098743
KEYWORDS Human papillomavirus
SOURCE Human papillomavirus
ORGANISM Human papillomavirus
Virusess; dsDNA viruses, no RNA stage; Papillomaviridae;
unclassified Papillomaviridae.
1 (bases 1 to 456)
Yamada,T., Wheeler,C.M., Halpern,A.L., Stewart,A.C., Hildesheim,A.
and Jenison,S.A.
Human papillomavirus type 16 variant lineages in United States
populations characterized by nucleotide sequence analysis of the
E6, E2, and L1 coding segments
J. Virol. 69 (12), 7743-7753 (1995)
7494284
2 (bases 1 to 456)
Farmer,A.D.
Direct Submission
Submitted (16-AUG-1995) Andrew D. Farmer, Los Alamos National
Laboratory, HIV Sequence Database, T-10, Mail Stop K710, Los
Alamos, NM 87501, USA
location/Qualifiers
1. .456
/organism="Human papillomavirus"
/mol_type="genomic DNA"
/isolate="OR 4541"
/db_xref="taxon:10566"
/note="collected by cervical lavage at Kaiser Permanente,
Portland, Oregon, United States."
1. .456
/gene="E6"
1. .456
/gene="E6"
/codon_start=1
/product="early transforming protein E6"
/protein_id="AA91666.1"
/db_xref="GI:1098744"
/translation="MFQDPQRPRIKLPDLCTELQTTIHDIILCEVCCKOQLRREYVD
FARPDLCIVRDGNPFAVCDKCKAFYSKISEYRYCYSYGTLEQQYNRPCLDLIR
CINQKPLCEBKQRHDKKQRFNIRGWTGRCMSCRCSRTRRETQL"

ORIGIN
Query Match 100.0%; Score 21; DB 10; Length 456;
Best Local Similarity 100.0%; Pred. No. 0.3; Mismatches 0; Indels 0; Gaps 0;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAGGACACAGTGGCTTTGAC 21
|||||
337 CAGGACACAGTGGCTTTGAC 317

RESULT 39

LOCUS	HPJ34120	456 bp	DNA	linear	VRL-08-MAR-1996
DEFINITION	Human papillomavirus type 16, isolate OR 4716, early transforming protein E6 (E6) gene, complete cds.				
ACCESSION	U34120				
VERSION	U34120.1 GI:1098745				
KEYWORDS					
SOURCE	Human papillomavirus				
ORGANISM	Human papillomavirus viruses; dsDNA viruses, no RNA stage; Papillomaviridae; unclassified Papillomaviridae. 1 (bases 1 to 456) Yamada,T., Wheeler,C.M., Halpern,A.L., Stewart,A.C., Hildesheim,A. and Jenson,S.A. Human papillomavirus type 16 variant lineages in United States populations characterized by nucleotide sequence analysis of the E6, E2, and L1 coding segments J. Virol. 69 (12), 7743-7753 (1995)				
TITLE	J. Virol. 69 (12), 7743-7753 (1995)				
JOURNAL	7494284				
PUBMED	2 (bases 1 to 456)				
REFERENCE	Farmer,A.D. Direct Submission Submitted (16-AUG-1995) Andrew D. Farmer, Los Alamos National Laboratory, HIV Sequence Database, T-10, Mail Stop K710, Los Alamos, NM 87501, USA				
AUTHORS	Location/Qualifiers				
TITLE	1..456				
JOURNAL	/organism="Human papillomavirus" /mol_type="genomic DNA" /isolate="OR 4716" /db_xref="taxon:10566" /note="collected by cervical lavage at Kaiser Permanente, Portland, Oregon, United States." 1..456 /gene="E6" 1..456 /gene="B6" /codon_start=1 /product="early transforming protein E6" /protein_id="AA91667.1" /db_xref="GI:1098746" /translation="MFDDPDRPRKLDPDLCTLTTHETHEILIECVCKQGLAREVDV PAFPLDCLVAEDGNPAVCDCIKLFEKISRYRVCYSIVGTLEQNGNLCOLLRLR CINCQKFLCEBKQRHLDKKQFRFNIRGRTGRCSRSSRTTRTQL"				
CDS					
gene					
ORIGIN	Query Match 100.0%; Score 21; DB 10; Length 456; Best Local Similarity 100.0%; Pred. No. 0.3; Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
OY	1 CAGGACACAGTGCGCTTTTGAC 21 				
Db	337 CAGGACACAGTGCGCTTTTGAC 317 				
RESULT 40					
HPJ34121/c					
LOCUS	HPJ34121 456 bp DNA linear VRL-08-MAR-1996				
DEFINITION	Human papillomavirus type 16, isolate OR 4724, early transforming protein E6 (E6) gene, complete cds.				
ACCESSION	U34121				
VERSION	U34121.1 GI:1098747				
KEYWORDS					
SOURCE	Human papillomavirus				
ORGANISM	Human papillomavirus viruses; dsDNA viruses, no RNA stage; Papillomaviridae; unclassified Papillomaviridae. 1 (bases 1 to 456) Yamada,T., Wheeler,C.M., Halpern,A.L., Stewart,A.C., Hildesheim,A. and Jenson,S.A. Human papillomavirus type 16 variant lineages in United States populations characterized by nucleotide sequence analysis of the E6, E2, and L1 coding segments				
REFERENCE					
AUTHORS					
TITLE					

```

JOURNAL PUBMED J Virol. 69 (12), 7743-7753 (1995)
REFERENCE 7494284
AUTHORS 2 (bases 1 to 456)
TITLE Farmer,A.D.
JOURNAL Direct Submission
Submitted (16-AUG-1995) Andrew D. Farmer, Los Alamos National
Laboratory, HIV Sequence Database, T-10, Mail Stop K710, Los
Alamos, NM 87501, USA
FEATURES
SOURCE Location/Qualifiers
1..456
/organism="Human papillomavirus"
/mol_type="genomic DNA"
/isolate="OR 4724"
/db_xref="taxon:10566"
/notes="collected by cervical lavage at Kaiser Permanente,
Portland, Oregon, United States."
1..456
/gene="E6"
1..456
/gene="E6"
1..456
/codon_start=1
/product="early transforming protein E6"
/protein_id="AA81668.1"
/db_xref="GI:1098748"
/translation="MEQDPQERPKLPLQCTELQTTIHDIILECVYCKQQLRREVD
FAFDFLCIVADGNPAVCDKCFYSKISBYSCVSLGTLLEQYNPLDLLLR
CINQKPLCEPKRQHIDKKQRFHINIGRWTCMSCRSSRTREIOL"
ORIGIN
Query Match 100.0%; Score 21; DB 10; Length 456;
Bee Local Similarity 100.0%; Pzed. No. 0.3;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CY 1 CAGGACACAGTGCTTTGAC 21
|||||
Db 337 CAGGACACAGTGCTTTGAC 317
|||||
RESULT 41
HPV34122 456 bp DNA linear VRL 08-MAR-1996
LOCUS Human papillomavirus type 16, isolate OR 4997, early transforming
DEFINITION protein E6 (E6) gene, complete cds.
ACCESSION U34122
VERSION U34122.1 GI:1098749
KEYWORDS
SOURCE Human papillomavirus
ORGANISM Human papillomavirus
Virusess: dsDNA viruses, no RNA stage; Papillomaviridae;
unclassified Papillomaviridae.
REFERENCE 1 (bases 1 to 456)
AUTHORS Yamada,T., Wheeler,C.M., Halpern,A.L., Stewart,A.C., Hildesheim,A.
and Jenson,S.A.
TITLE Human papillomavirus type 16 variant lineages in United States
populations characterized by nucleotide sequence analysis of the
E6, E2, and E1 coding segments
JOURNAL J. Virol. 69 (12), 7743-7753 (1995)
PUBMED 7494284
JOURNAL 2 (bases 1 to 456)
REFERENCE Farmer,A.D.
AUTHORS Direct Submission
TITLE Submitted (16-AUG-1995) Andrew D. Farmer, Los Alamos National
JOURNAL Laboratory, HIV Sequence Database, T-10, Mail Stop K710, Los
Alamos, NM 87501, USA
FEATURES
SOURCE Location/Qualifiers
1..456
/organism="Human papillomavirus"
/mol_type="genomic DNA"
/isolate="OR 4997"
/db_xref="taxon:10566"
/notes="collected by cervical lavage at Kaiser Permanente,
Portland, Oregon, United States."
1..456

```

CDS

/gene="E6"
1. .456
/gene="E6"
/codon_start=1
/product="early transforming protein E6"
/protein_id="AA91669.1"
/db_xref="GI:1098750"
/translation="MFODPQERPRKLPOLCTELQTTIHILBCVCKQQLRREYVD
FARPDLCIVRDGNPYAVCDCKFKYSKISEYHYCYSLYGTLEQQYNPCLDILLR
CINQOKPLCEBKQRHDKQRFHNIRGWTGRCMSCRSRTRRETOL"

ORIGIN

Query Match 100.0%; Score 21; DB 10; Length 456;
Best Local Similarity 100.0%; Pred. No. 0.3;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAGGACACAGTGGCTTTGAC 21
337 CAGGACACAGTGGCTTTGAC 317

RESULT 42
HPJ34123/c 456 bp DNA linear VRL 08-MAR-1996
LOCUS
DEFINITION Human papillomavirus type 16, isolate OR 5110, early transforming
protein E6 (E6) gene, complete cds.
ACCESSION U34123.1 GI:1098751
VERSION
KEYWORDS
SOURCE Human papillomavirus
ORGANISM Human papillomavirus
VIRUSES; dsDNA viruses, no RNA stage; Papillomaviridae;
unclassified Papillomaviridae.
REFERENCE 1 (bases 1 to 456)
AUTHORS Yamada,T., Wheeler,C.M., Halpern,A.L., Stewart,A.C., Hildesheim,A.
and Jenison,S.A.
TITLE Human papillomavirus type 16 variant lineages in United States
populations characterized by nucleotide sequence analysis of the
E6, E2, and L1 coding segments
J. Virol. 69 (12), 7743-7753 (1995)
7494284
2 (bases 1 to 456)
REFERENCE PubMed
AUTHORS Farmer,A.D.
TITLE Direct Submission
JOURNAL Submitted (16-AUG-1995) Andrew D. Farmer, Los Alamos National
Laboratory, HIV Sequence Database, T-10, Mail Stop K710, Los
Alamos, NM 87501, USA
location/Qualifiers
1. .456
/organism="Human papillomavirus"
/mol_type="genomic DNA"
/isolate="OR 5110"
/db_xref="taxon:10566"
/note="collected by cervical lavage at Kaiser Permanente,
Portland, Oregon, United States."
1. .456
/gene="E6"
1. .456
/gene="E6"
/codon_start=1
/product="early transforming protein E6"
/protein_id="AA91670.1"
/db_xref="GI:1098752"
/translation="MFODPQERPRKLPOLCTELQTTIHILBCVCKQQLRREYVD
FARPDLCIVRDGNPYAVCDCKFKYSKISEYHYCYSLYGTLEQQYNPCLDILLR
CINQOKPLCEBKQRHDKQRFHNIRGWTGRCMSCRSRTRRETOL"

ORIGIN

Query Match 100.0%; Score 21; DB 10; Length 456;
Best Local Similarity 100.0%; Pred. No. 0.3;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAGGACACAGTGGCTTTGAC 21
1 CAGGACACAGTGGCTTTGAC 21

Db 337 CAGGACACAGTGGCTTTGAC 317

RESULT 43
HPJ34124/c 456 bp DNA linear VRL 08-MAR-1996
LOCUS
DEFINITION Human papillomavirus type 16, isolate OR 5428, early transforming
protein E6 (E6) gene, complete cds.
ACCESSION U34124.1 GI:1098753
VERSION
KEYWORDS
SOURCE Human papillomavirus
ORGANISM Human papillomavirus
VIRUSES; dsDNA viruses, no RNA stage; Papillomaviridae;
unclassified Papillomaviridae.
REFERENCE 1 (bases 1 to 456)
AUTHORS Yamada,T., Wheeler,C.M., Halpern,A.L., Stewart,A.C., Hildesheim,A.
and Jenison,S.A.
TITLE Human papillomavirus type 16 variant lineages in United States
populations characterized by nucleotide sequence analysis of the
E6, E2, and L1 coding segments
J. Virol. 69 (12), 7743-7753 (1995)
7494284
2 (bases 1 to 456)
REFERENCE PubMed
AUTHORS Farmer,A.D.
TITLE Direct Submission
JOURNAL Submitted (16-AUG-1995) Andrew D. Farmer, Los Alamos National
Laboratory, HIV Sequence Database, T-10, Mail Stop K710, Los
Alamos, NM 87501, USA
location/Qualifiers
1. .456
/organism="Human papillomavirus"
/mol_type="genomic DNA"
/isolate="OR 5428"
/db_xref="taxon:10566"
/note="collected by cervical lavage at Kaiser Permanente,
Portland, Oregon, United States."
1. .456
/gene="E6"
1. .456
/gene="E6"
/codon_start=1
/product="early transforming protein E6"
/protein_id="AA91671.1"
/db_xref="GI:1098754"
/translation="MFODPQERPRKLPOLCTELQTTIHILBCVCKQQLRREYVD
FARPDLCIVRDGNPYAVCDCKFKYSKISEYHYCYSLYGTLEQQYNPCLDILLR
CINQOKPLCEBKQRHDKQRFHNIRGWTGRCMSCRSRTRRETOL"

ORIGIN

Query Match 100.0%; Score 21; DB 10; Length 456;
Best Local Similarity 100.0%; Pred. No. 0.3;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAGGACACAGTGGCTTTGAC 21
337 CAGGACACAGTGGCTTTGAC 317

RESULT 44
HPJ34125/c 456 bp DNA linear VRL 08-MAR-1996
LOCUS
DEFINITION Human papillomavirus type 16, isolate OR 6106, early transforming
protein E6 (E6) gene, complete cds.
ACCESSION U34125.1 GI:1098755
VERSION
KEYWORDS
SOURCE Human papillomavirus
ORGANISM Human papillomavirus
VIRUSES; dsDNA viruses, no RNA stage; Papillomaviridae;
unclassified Papillomaviridae.
REFERENCE 1 (bases 1 to 456)

AUTHORS Yamada,T., Wheeler,C.M., Halpern,A.L., Stewart,A.C., Hildesheim,A.
and Jenison,S.A.
TITLE Human papillomavirus type 16 variant lineages in United States
populations characterized by nucleotide sequence analysis of the
E6, E2, and L1 coding segments
J. Virol. 69 (12), 7743-7753 (1995)
JOURNAL PUBMED 7494284
REFERENCE 2 (bases 1 to 456)
AUTHORS Farmer,A.D.
TITLE Direct Submision
Submitted (16-AUG-1995) Andrew D. Farmer, Los Alamos National
Laboratory, HIV Sequence Database, T-10, Mail Stop K710, Los
Alamos, NM 87501, USA
JOURNAL location/Qualifiers
1. .456
source /organism="Human papillomavirus"
/mol_type="genomic DNA"
/isolate="OR 6106"
/db_xref="taxon:10566"
/note="collected by cervical lavage at Kaiser Permanente,
Portland, Oregon, United States."
1. .456
/gene="E6"
1. .456
/codon_start=1
/product="early transforming protein E6"
/protein_id="AA91672.1"
/db_xref="GI:1098756"
/translation="MFQDPQERPRKLPOLCTELQTTIHDIIECVYCKQQLRREYVD
FAFRDLCIVRDGNPYAVCDKCLKFYSKISRYHYCVSYGTTLEGOYKPLCDLLIR
CINQKPLCPBEKQRLDKQRFHNRGWTGRCMSCSSRRTRRTQ"

ORIGIN
Query Match 100.0%; Score 21; DB 10; Length 456;
Best Local Similarity 100.0%; Pred. No. 0.3;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAGGACACAGTGCGCTTTGAC 21
Db 337 CAGGACACAGTGCGCTTTGAC 317

RESULT 45
HPV34126/c 456 bp DNA linear VRL 08-MAR-1996
LOCUS Human papillomavirus type 16, isolate OR 6170, early transforming
DEFINITION protein E6 (E6) gene, complete cds.
ACCESSION U34126
VERSION U34126.1 GI:1098757
KEYWORDS
SOURCE Human papillomavirus
ORGANISM Human papillomavirus
VIRUSES; deDNA viruses, no RNA stage; Papillomaviridae;
unclassified Papillomaviridae.
1 (bases 1 to 456)
REFERENCE Yamada,T., Wheeler,C.M., Halpern,A.L., Stewart,A.C., Hildesheim,A.
and Jenison,S.A.
TITLE Human papillomavirus type 16 variant lineages in United States
populations characterized by nucleotide sequence analysis of the
E6, E2, and L1 coding segments
J. Virol. 69 (12), 7743-7753 (1995)
JOURNAL PUBMED 7494284
REFERENCE 2 (bases 1 to 456)
AUTHORS Farmer,A.D.
TITLE Direct Submision
Submitted (16-AUG-1995) Andrew D. Farmer, Los Alamos National
Laboratory, HIV Sequence Database, T-10, Mail Stop K710, Los
Alamos, NM 87501, USA
JOURNAL location/Qualifiers
1. .456
source /organism="Human papillomavirus"
/mol_type="genomic DNA"

gene /isolate="OR 6170"
/db_xref="taxon:10566"
/note="collected by cervical lavage at Kaiser Permanente,
Portland, Oregon, United States."
1. .456
/gene="E6"
1. .456
/codon_start=1
/product="early transforming protein E6"
/protein_id="AA91673.1"
/db_xref="GI:1098758"
/translation="MFQDPQERPRKLPOLCTELQTTIHDIIECVYCKQQLRREYVD
FAFRDLCIVRDGNPYAVCDKCLKFYSKISRYHYCVSYGTTLEGOYKPLCDLLIR
CINQKPLCPBEKQRLDKQRFHNRGWTGRCMSCSSRRTRRTQ"

ORIGIN
Query Match 100.0%; Score 21; DB 10; Length 456;
Best Local Similarity 100.0%; Pred. No. 0.3;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAGGACACAGTGCGCTTTGAC 21
Db 337 CAGGACACAGTGCGCTTTGAC 317

RESULT 46
HPV34127/c 456 bp DNA linear VRL 08-MAR-1996
LOCUS Human papillomavirus type 16, isolate OR 6311, early transforming
DEFINITION protein E6 (E6) gene, complete cds.
ACCESSION U34127
VERSION U34127.1 GI:1098759
KEYWORDS
SOURCE Human papillomavirus
ORGANISM Human papillomavirus
VIRUSES; deDNA viruses, no RNA stage; Papillomaviridae;
unclassified Papillomaviridae.
1 (bases 1 to 456)
REFERENCE Yamada,T., Wheeler,C.M., Halpern,A.L., Stewart,A.C., Hildesheim,A.
and Jenison,S.A.
TITLE Human papillomavirus type 16 variant lineages in United States
populations characterized by nucleotide sequence analysis of the
E6, E2, and L1 coding segments
J. Virol. 69 (12), 7743-7753 (1995)
JOURNAL PUBMED 7494284
REFERENCE 2 (bases 1 to 456)
AUTHORS Farmer,A.D.
TITLE Direct Submision
Submitted (16-AUG-1995) Andrew D. Farmer, Los Alamos National
Laboratory, HIV Sequence Database, T-10, Mail Stop K710, Los
Alamos, NM 87501, USA
JOURNAL location/Qualifiers
1. .456
source /organism="Human papillomavirus"
/mol_type="genomic DNA"
/isolate="OR 6311"
/db_xref="taxon:10566"
/note="collected by cervical lavage at Kaiser Permanente,
Portland, Oregon, United States."
1. .456
/gene="E6"
1. .456
/codon_start=1
/product="early transforming protein E6"
/protein_id="AA91674.1"
/db_xref="GI:1098760"
/translation="MFQDPQERPRKLPOLCTELQTTIHDIIECVYCKQQLRREYVD
FAFRDLCIVRDGNPYAVCDKCLKFYSKISRYHYCVSYGTTLEGOYKPLCDLLIR
CINQKPLCPBEKQRLDKQRFHNRGWTGRCMSCSSRRTRRTQ"

Query Match 100.0%; Score 21; DB 10; Length 456;
Best Local Similarity 100.0%; Pred. No. 0.3;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CAGGACACAGTGGCTTTTGAC 21
Db 337 CAGGACACAGTGGCTTTTGAC 317

RESULT 47
HPU34128/c 456 bp DNA linear VRL 08-MAR-1996
LOCUS HPU34128
DEFINITION Human papillomavirus type 16, isolate OR 7145, early transforming protein E6 (E6) gene, complete cds.
ACCESSION U34128
VERSION U34128.1 GI:1098761
KEYWORDS
SOURCE Human papillomavirus
ORGANISM Human papillomavirus
VIRUSES; dsDNA viruses, no RNA stage; Papillomaviridae;
REFERENCE
AUTHORS Yamada,T., Wheeler,C.M., Halpern,A.L., Stewart,A.C., Hildeheim,A. and Jenison,S.A.
TITLE Human papillomavirus type 16 variant lineages in United States populations characterized by nucleotide sequence analysis of the E6, E2, and L1 coding segments
JOURNAL J. Virol. 69 (12), 7743-7753 (1995)
PUBMED 7494284
REFERENCE 2 (bases 1 to 456)
AUTHORS Farmer,A.D.
TITLE Direct Submission
JOURNAL Submitted (16-AUG-1995) Andrew D. Farmer, Los Alamos National Laboratory, HIV Sequence Database, T-10, Mail Stop K710, Los Alamos, NM 87501, USA
FEATURES
source
1. .456
/organism="Human papillomavirus"
/mol_type="genomic DNA"
/isolate="OR 7145"
/db_xref="taxon:10566"
/note="collected by cervical lavage at Kaiser Permanente, Portland, Oregon, United States."
1. .456
/gene="E6"
1. .456
/gene="E6"
/codon_start=1
/product="early transforming protein E6"
/protein_id="AA91675.1"
/db_xref="GI:1098762"
/translation="MFQDPQEPRIKLPDLTEIQTTHIHILGCVCKQQLRREYD FARRDLCIVRDGNPVAVCDCIKFYSKISEYRYCSLYGTTLEQYNPLCDLLIR CINQKPLCPBEKQRIHDKQRFNIRGWTGRCMSCRSRRTEROL"

ORIGIN
Query Match 100.0%; Score 21; DB 10; Length 456;
Best Local Similarity 100.0%; Pred. No. 0.3;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CAGGACACAGTGGCTTTTGAC 21
Db 337 CAGGACACAGTGGCTTTTGAC 317

RESULT 48
HPU34129/c 456 bp DNA linear VRL 08-MAR-1996
LOCUS HPU34129
DEFINITION Human papillomavirus type 16, isolate OR 7574, early transforming protein E6 (E6) gene, complete cds.
ACCESSION U34129
VERSION U34129.1 GI:1098763
KEYWORDS

SOURCE Human papillomavirus
ORGANISM Human papillomavirus
VIRUSES; dsDNA viruses, no RNA stage; Papillomaviridae;
REFERENCE
AUTHORS Yamada,T., Wheeler,C.M., Halpern,A.L., Stewart,A.C., Hildeheim,A. and Jenison,S.A.
TITLE Human papillomavirus type 16 variant lineages in United States populations characterized by nucleotide sequence analysis of the E6, E2, and L1 coding segments
JOURNAL J. Virol. 69 (12), 7743-7753 (1995)
PUBMED 7494284
REFERENCE 2 (bases 1 to 456)
AUTHORS Farmer,A.D.
TITLE Direct Submission
JOURNAL Submitted (16-AUG-1995) Andrew D. Farmer, Los Alamos National Laboratory, HIV Sequence Database, T-10, Mail Stop K710, Los Alamos, NM 87501, USA
FEATURES
source
1. .456
/organism="Human papillomavirus"
/mol_type="genomic DNA"
/isolate="OR 7574"
/db_xref="taxon:10566"
/note="collected by cervical lavage at Kaiser Permanente, Portland, Oregon, United States."
1. .456
/gene="E6"
1. .456
/gene="E6"
/codon_start=1
/product="early transforming protein E6"
/protein_id="AA91676.1"
/db_xref="GI:1098764"
/translation="MFQDPQEPRIKLPDLTEIQTTHIHILGCVCKQQLRREYD FARRDLCIVRDGNPVAVCDCIKFYSKISEYRYCSLYGTTLEQYNPLCDLLIR CINQKPLCPBEKQRIHDKQRFNIRGWTGRCMSCRSRRTEROL"

ORIGIN
Query Match 100.0%; Score 21; DB 10; Length 456;
Best Local Similarity 100.0%; Pred. No. 0.3;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CAGGACACAGTGGCTTTTGAC 21
Db 337 CAGGACACAGTGGCTTTTGAC 317

RESULT 49
HPU34130/c 456 bp DNA linear VRL 08-MAR-1996
LOCUS HPU34130
DEFINITION Human papillomavirus type 16, isolate OR 7587, early transforming protein E6 (E6) gene, complete cds.
ACCESSION U34130
VERSION U34130.1 GI:1098765
KEYWORDS
SOURCE Human papillomavirus
ORGANISM Human papillomavirus
VIRUSES; dsDNA viruses, no RNA stage; Papillomaviridae;
REFERENCE
AUTHORS Yamada,T., Wheeler,C.M., Halpern,A.L., Stewart,A.C., Hildeheim,A. and Jenison,S.A.
TITLE Human papillomavirus type 16 variant lineages in United States populations characterized by nucleotide sequence analysis of the E6, E2, and L1 coding segments
JOURNAL J. Virol. 69 (12), 7743-7753 (1995)
PUBMED 7494284
REFERENCE 2 (bases 1 to 456)
AUTHORS Farmer,A.D.
TITLE Direct Submission
JOURNAL Submitted (16-AUG-1995) Andrew D. Farmer, Los Alamos National Laboratory, HIV Sequence Database, T-10, Mail Stop K710, Los

FEATURES Alamos, NM 87501, USA
Location/Qualifiers
1..456
/organism="Human papillomavirus"
/mol_type="genomic DNA"
/isolate="OR 7632"
/db_xref="taxon:10566"
/note="collected by cervical lavage at Kaiser Permanente, Portland, Oregon, United States."
1..456
/gene="E6"
1..456
/gene="E6"
/codon_start=1
/product="early transforming protein E6"
/protein_id="AA91677.1"
/db_xref="GI:1098766"
/translation="MFQDPQERPTKLPDLCTELQTTIHDIILFCVYCKQQLRREYDFAFRDLCIVRDGNPYAVCDKCLKFYSKISEYRYCYSLYGTLEQQYNKPLCDLLIR
CINCKPLCPBEKQRHLDKKQRFHNRGHWGRCMCCRSSRTRRETOL"

ORIGIN

Query Match 100.0%; Score 21; DB 10; Length 456;
Best Local Similarity 100.0%; Pred. No. 0.3;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAGACACAGTGGCTTTTGAC 21
|||||
Db 337 CAGACACAGTGGCTTTTGAC 317

RESULT 50
HPV34131/c 456 bp DNA linear VRL 08-MAR-1996
LOCUS Human papillomavirus type 16, isolate OR 7632, early transforming
DEFINITION protein E6 (E6) gene, complete cds.
ACCESSION U34131
VERSION U34131.1 GI:1098767
KEYWORDS
SOURCE Human papillomavirus
ORGANISM Human papillomavirus
Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
unclassified Papillomaviridae.
1 (bases 1 to 456)
Yamada, T., Wheeler, C.M., Halpern, A.L., Stewart, A.C., Hildesheim, A.
and Jensen, S.A.
Human papillomavirus type 16 variant lineages in United States
populations characterized by nucleotide sequence analysis of the
E6, E2, and L1 coding segments
J. Virol. 69 (12), 7743-7753 (1995)
7494284
2 (bases 1 to 456)
Farmer, A.D.
Direct Submission
Submitted (16-AUG-1995) Andrew D. Farmer, Los Alamos National
Laboratory, HIV Sequence Database, T-10, Mail Stop K710, Los
Alamos, NM 87501, USA
Location/Qualifiers
1..456
/organism="Human papillomavirus"
/mol_type="genomic DNA"
/isolate="OR 7632"
/db_xref="taxon:10566"
/note="collected by cervical lavage at Kaiser Permanente,
Portland, Oregon, United States."
1..456
/gene="E6"
1..456
/gene="E6"
/codon_start=1
/product="early transforming protein E6"
/protein_id="AA91678.1"
/db_xref="GI:1098768"

gene
CDS

/translation="MFQDPQERPTKLPDLCTELQTTIHDIILFCVYCKQQLRREYDFAFRDLCIVRDGNPYAVCDKCLKFYSKISEYRYCYSLYGTLEQQYNKPLCDLLIR
CINCKPLCPBEKQRHLDKKQRFHNRGHWGRCMCCRSSRTRRETOL"

ORIGIN

Query Match 100.0%; Score 21; DB 10; Length 456;
Best Local Similarity 100.0%; Pred. No. 0.3;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAGACACAGTGGCTTTTGAC 21
|||||
Db 337 CAGACACAGTGGCTTTTGAC 317

Search completed: May 24, 2006, 06:54:50
Job time : 698.284 secs

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 24, 2006, 05:46:37 ; Search time 972.624 Seconds
(without alignments)
2103.915 Million cell updates/sec

Title: US-10-601-913-121

Perfect score: 32

Sequence: 1 TTATTAATAGGCGCTGCGGTGCCGAAC 32

Scoring table: IDENTITY_NUC

Searched: 6366136 seqs, 31973710525 residues

Total number of hits satisfying chosen parameters: 12732272

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 1000 summaries

Database : GenBml:*
1: gb env:*
2: gb_pat:*
3: gb_ph:*
4: gb_pl:*
5: gb_pr:*
6: gb_ro:*
7: gb_sts:*
8: gb_ay:*
9: gb_un:*
10: gb_vl:*
11: gb_ov:*
12: gb_htg:*
13: gb_in:*
14: gb_om:*
15: gb_da:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match Length	DB ID	Description
1	32	100.0	32	2	AR349181 Sequence
2	32	100.0	32	2	AR349182 Sequence
3	32	100.0	32	2	AR349183 Sequence
4	32	100.0	32	2	AR349184 Sequence
5	32	100.0	476	2	AR080321 Sequence
6	32	100.0	476	2	AR085249 Sequence
7	32	100.0	477	2	CS113238 Sequence
8	32	100.0	528	10	PAHPV181
9	32	100.0	817	2	188814
10	32	100.0	837	2	A98759 Sequence 20
11	32	100.0	837	2	BD080320 Vaccine.
12	32	100.0	837	2	BD103169 Sequence.
13	32	100.0	837	2	AR183581 Sequence.
14	32	100.0	837	2	AX020943 Sequence
15	32	100.0	843	10	HPV18491
16	32	100.0	912	10	PPH18A
17	32	100.0	1000	2	A06328
18	32	100.0	1000	2	A07621 Amplified s

19	32	100.0	1152	2	A98761 Sequence 22
20	32	100.0	1152	2	BD080321 Vaccine.
21	32	100.0	1152	2	BD103170 Sequence.
22	32	100.0	1152	2	AR183582 Sequence
23	32	100.0	1152	2	AX020945 Sequence
24	32	100.0	1152	2	BD015779 Method fo
25	32	100.0	1188	2	ES4157 Method for
26	32	100.0	1188	10	X04354 Human papil
27	32	100.0	1750	2	A06324 Hpv18 genes
28	32	100.0	1750	10	PAHPV18E
29	32	100.0	1315	10	HUMHELB
30	32	100.0	5210	10	HPU89349
31	32	100.0	7857	2	CS073239
32	32	100.0	7857	10	AY262282
33	32	100.0	7857	10	PAHPV18
34	30.4	95.0	910	10	HPV45BCE7
35	30.4	95.0	1843	10	PPH45E67A
36	30.4	95.0	7858	10	HPV45
37	30.4	95.0	8039	5	HS242956
38	24.6	76.9	866	6	MUSRP
39	24.2	75.6	1225	5	BC056907
40	24.2	75.6	1300	10	HUMHELA
41	24	75.0	1095	2	AR084350
42	24	75.0	1108	2	AR084344
43	24	75.0	1110	2	AR084343
44	24	75.0	1110	2	AR084345
45	24	75.0	1313	10	HPV46
46	24	75.0	3283	2	AR084352
47	24	75.0	3283	10	HPU22461
48	24	75.0	4618	10	HSPFUSION
49	24	75.0	6042	10	HUMHPVME18
50	24	75.0	7822	10	DO080079
51	24	75.0	7833	2	160253
52	24	75.0	7833	2	164482
53	24	75.0	7833	10	PPH39
54	24	75.0	7905	2	CS073246
55	24	75.0	7905	10	HPU21941
56	24	74.4	630	2	AX400844
57	23.8	74.4	778	6	RATRP17
58	23.8	74.4	836	11	EX933673
59	23.8	74.4	836	11	EX935217
60	23.8	74.4	836	11	EX936251
61	23.8	74.4	885	2	AX552259
62	23.8	74.4	221962	12	AC136580
63	23.6	73.8	16099	5	AL589786
64	23.6	73.8	161632	12	AC053508
65	23.6	73.8	163118	12	AL355528
66	23.6	73.8	168281	5	AL450108
67	23	71.9	30	2	AR077014
68	23	71.9	30	2	AR099824
69	23	71.9	30	2	AR142235
70	23	71.9	30	2	134496
71	23	71.9	30	2	157331
72	23	71.9	30	2	173209
73	23	71.9	7759	10	HPU37488
74	23	71.9	7812	10	AP131950
75	23	71.9	94092	12	AC155398
76	22.6	70.6	167	2	CQ476121
77	22.4	70.0	137036	15	AJ698720
78	22.4	70.0	172732	6	AC113039
79	22.4	70.0	174313	12	AC165234
80	22.4	70.0	207067	7	AC161108
81	22.2	69.4	201	7	BV198704
82	22.2	69.4	207	2	CQ715203
83	22.2	69.4	460	2	CQ049241
84	22.2	69.4	460	2	CQ064262
85	22.2	69.4	460	2	CQ091202
86	22.2	69.4	460	2	CQ130040
87	22.2	69.4	460	2	CQ168659
88	22.2	69.4	460	2	CQ197782
89	22.2	69.4	460	2	CQ213224
90	22.2	69.4	460	2	CQ251803
91	22.2	69.4	460	2	CQ288957

92	22.2	69.4	460	2	CQ325963	Sequence	c 165	22	68.8	51	2	AX803129
93	22.2	69.4	581	2	CS211030	Sequence	166	22	68.8	650	2	C0671722
94	22.2	69.4	581	2	CS214672	Sequence	c 167	22	68.8	13338	12	AC153776
95	22.2	69.4	659	2	AX552261	Sequence	168	21.6	67.5	120782	12	AC144505
96	22.2	69.4	741	2	CQ737742	Sequence	169	21.4	66.9	582	10	G77890
97	22.2	69.4	750	2	CQ054321	Sequence	170	21.4	66.9	7717	7	AF436129
98	22.2	69.4	750	2	CQ073558	Sequence	c 171	21.4	66.9	180784	12	AC160487
99	22.2	69.4	750	2	CQ104440	Sequence	c 172	21.4	66.9	186611	6	AC119215
100	22.2	69.4	750	2	CQ143159	Sequence	c 173	21.4	66.9	193136	12	AC095585
101	22.2	69.4	750	2	CQ178644	Sequence	174	21.4	66.9	220936	12	AC164869
102	22.2	69.4	750	2	CQ020993	Sequence	c 175	21.4	66.9	222448	12	AC161085
103	22.2	69.4	750	2	CQ226340	Sequence	176	21.4	66.9	246551	12	AC096478
104	22.2	69.4	750	2	CQ264482	Sequence	c 177	21.4	66.9	249620	6	AC161484
105	22.2	69.4	750	2	CQ301577	Sequence	c 178	21.4	66.9	252524	12	AC114081
106	22.2	69.4	824	2	CQ338792	Sequence	179	21.4	66.9	269435	12	AC167349
107	22.2	69.4	845	5	CQ716065	Sequence	c 180	21.4	66.9	269435	12	AC167349
108	22.2	69.4	845	5	AB169738	Macaca fa	181	21.2	66.2	123	2	CQ080989
109	22.2	69.4	1023	2	CQ490614	Sequence	182	21.2	66.2	123	2	CQ115474
110	22.2	69.4	1023	2	CQ493111	Sequence	183	21.2	66.2	123	2	CQ154282
111	22.2	69.4	1023	2	CQ493249	Sequence	184	21.2	66.2	123	2	CQ186859
112	22.2	69.4	1023	2	CQ493345	Sequence	185	21.2	66.2	123	2	CQ237529
113	22.2	69.4	1023	2	CQ493672	Sequence	186	21.2	66.2	123	2	CQ275152
114	22.2	69.4	1023	2	CQ493995	Sequence	187	21.2	66.2	123	2	CQ312126
115	22.2	69.4	1023	2	CQ496449	Sequence	188	21.2	66.2	123	2	CQ349522
116	22.2	69.4	1033	14	BT021516	BT021516 Bos tauru	189	21.2	66.2	354	2	CQ734529
117	22.2	69.4	1033	5	AC126748	AC126748 Homo sapi	190	21.2	66.2	380	2	CQ676372
118	22.2	69.4	1033	5	AC158317	AC158317 Homo sapi	191	21.2	66.2	598	2	CQ071784
119	22.2	69.4	38703	5	AC000031	AC000031 Homo sapi	192	21.2	66.2	598	2	CQ102405
120	22.2	69.4	44422	12	AC146440	AC146440 Pan trogl	193	21.2	66.2	598	2	CQ114332
121	22.2	69.4	99766	12	AL153702	AL153702 Homo sapi	194	21.2	66.2	598	2	CQ116982
122	22.2	69.4	103186	5	AL158218	AL158218 Homo sapi	195	21.2	66.2	598	2	CQ224613
123	22.2	69.4	106735	12	AC101381	AC101381 Homo sapi	196	21.2	66.2	598	2	CQ262631
124	22.2	69.4	110000	12	AC106662_1	Continuation (2 of	197	21.2	66.2	598	2	CQ299688
125	22.2	69.4	110000	12	AC106662_2	Continuation (3 of	198	21.2	66.2	598	2	CQ336842
126	22.2	69.4	112911	5	AL355310	AL355310 Human DNA	199	21.2	66.2	123165	5	AC008771
127	22.2	69.4	126807	5	HS391022	AL031577 Human DNA	c 200	21.2	66.2	126052	5	AC018764
128	22.2	69.4	132764	5	AC073842	AC073842 Homo sapi	c 201	21.2	66.2	130117	5	AC004907
129	22.2	69.4	145224	12	AC009856	AC009856 Homo sapi	c 202	21.2	66.2	167664	12	AC027303
130	22.2	69.4	146224	5	AC005883	AC005883 Homo sapi	c 203	21.2	66.2	181226	5	AC012312
131	22.2	69.4	160439	5	AC005345	AC005345 Homo sapi	c 204	21.2	66.2	184598	5	AL139406
132	22.2	69.4	163332	5	AC004821	AC004821 Homo sapi	c 205	21.2	66.2	194874	12	AC080090
133	22.2	69.4	164702	2	AX706960	AX706960 Sequence	c 206	21.2	66.2	246164	12	AC098751
134	22.2	69.4	164702	2	AX707890	AX707890 Sequence	c 207	21.2	66.2	253402	5	AC008534
135	22.2	69.4	167159	2	AC113403	AC113403 Homo sapi	c 208	21.2	66.2	263827	12	AC094579
136	22.2	69.4	167854	12	AC092483	AC092483 Homo sapi	c 209	21.2	66.2	322972	12	AC129853
137	22.2	69.4	170819	12	AC091780	AC091780 Homo sapi	c 210	21	65.6	420	2	AR349169
138	22.2	69.4	172043	12	AC055750	AC055750 Homo sapi	c 211	21	65.6	22	2	AR349170
139	22.2	69.4	175281	5	AC092859	AC092859 Pan trogl	c 212	21	65.6	22	2	AR349171
140	22.2	69.4	175794	5	AC147031	AC147031 Pan trogl	c 213	21	65.6	22	2	AR349172
141	22.2	69.4	179583	12	AC011931	AC011931 Homo sapi	c 214	21	65.6	420	2	CQ471329
142	22.2	69.4	179616	5	AC078994	AC078994 Homo sapi	c 215	21	65.6	55677	12	AC154949
143	22.2	69.4	182256	5	AC005058	AC005058 Homo sapi	c 216	21	65.6	110000	4	AP008207
144	22.2	69.4	183155	5	AC073021	AC073021 Homo sapi	c 217	21	65.6	151630	12	AC169465
145	22.2	69.4	188167	12	AC068983	AC068983 Homo sapi	c 218	21	65.6	171849	12	AC175179
146	22.2	69.4	189863	12	AC027463	AC027463 Homo sapi	c 219	21	65.6	181997	12	AC165359
147	22.2	69.4	193986	12	AC130783	AC130783 Pan trogl	c 220	21	65.6	182709	12	AC148374
148	22.2	69.4	197252	5	AC170498	AC170498 Bos tauru	c 221	21	65.6	188707	12	AC162039
149	22.2	69.4	197360	2	AC117401	AC117401 Homo sapi	c 222	21	65.6	190659	12	AC170342
150	22.2	69.4	200536	5	AL356093	AL356093 Human DNA	c 223	21	65.6	191022	4	AP003431
151	22.2	69.4	208310	12	AC166382	AC166382 Bos tauru	c 224	21	65.6	222962	12	AC160677
152	22.2	69.4	213601	12	AC165037	AC165037 Bos tauru	c 225	21	65.6	230115	12	AC095927
153	22.2	69.4	224862	12	AC125660	AC125660 Rattus no	c 226	21	65.6	237025	12	AC149690
154	22.2	69.4	229155	5	AC018755	AC018755 Homo sapi	c 227	20.8	65.0	757	6	MUSRPL7R
155	22.2	69.4	237165	6	AC167216	AC167216 Bos tauru	c 228	20.8	65.0	899	6	BC096452
156	22.2	69.4	237532	12	AC094950	AC094950 Rattus no	c 229	20.8	65.0	906	6	BC025909
157	22.2	69.4	240792	12	AC130623	AC130623 Rattus no	c 230	20.8	65.0	915	6	BC051261
158	22.2	69.4	247475	12	AC008247	AC008247 Homo sapi	c 231	20.8	65.0	962	6	BC086786
159	22.2	69.4	251835	12	AC157033	AC157033 Bos tauru	c 232	20.8	65.0	1867	5	HSU80579
160	22.2	69.4	258216	12	AC156331	AC156331 Bos tauru	c 233	20.8	65.0	2455	11	XLU63920
161	22.2	69.4	275862	12	AC112376	AC112376 Rattus no	c 234	20.8	65.0	2639	11	BC045089
162	22.2	69.4	284106	12	AC166087	AC166087 Bos tauru	c 235	20.8	65.0	3112	11	XLU63919
163	22.2	68.8	51	2	DD178844	DD178844 METHOD FO	c 236	20.8	65.0	3418	6	MUSRPL7A
164	22.2	68.8	51	2	AX802744	AX802744 Sequence	c 237	20.8	65.0	7871	10	AB027021

238	20.8	65.0	7904	10	AF293961	311	20.6	64.4	471	2	C0696169	Sequence
C 239	20.8	65.0	75296	5	AC110795	312	20.6	64.4	473	2	C0679633	Sequence
C 240	20.8	65.0	83518	12	AC117971	313	20.6	64.4	479	2	C0687119	Sequence
241	20.8	65.0	153292	5	AL51316	314	20.6	64.4	490	2	C0711883	Sequence
242	20.8	65.0	160008	6	AC158217	315	20.6	64.4	492	2	AX360681	Sequence
C 243	20.8	65.0	161949	6	CT485613	316	20.6	64.4	498	2	C0712880	Sequence
C 244	20.8	65.0	169321	6	AC107725	317	20.6	64.4	500	2	C0712891	Sequence
245	20.8	65.0	171102	12	CT476835	318	20.6	64.4	513	2	C0670105	Sequence
C 246	20.8	65.0	171864	12	AC016248	319	20.6	64.4	514	2	C0672486	Sequence
247	20.8	65.0	174615	12	AC021249	320	20.6	64.4	548	2	C0832487	Sequence
248	20.8	65.0	176870	6	AL844583	321	20.6	64.4	556	2	C0671702	Sequence
249	20.8	65.0	184674	12	AC102557	322	20.6	64.4	567	2	C0669223	Sequence
250	20.8	65.0	185440	12	AC015459	323	20.6	64.4	572	2	C0673737	Sequence
251	20.8	65.0	189335	6	AC112945	324	20.6	64.4	584	2	C0700469	Sequence
252	20.8	65.0	190466	6	AC163446	325	20.6	64.4	591	2	C0832470	Sequence
C 253	20.8	65.0	201383	6	AC162442	326	20.6	64.4	591	2	C066967	Sequence
C 254	20.8	65.0	202155	6	AC113269	327	20.6	64.4	596	7	BV250070	Sequence
C 255	20.8	65.0	216945	12	AC156619	328	20.6	64.4	605	2	C0671503	Sequence
C 256	20.8	65.0	218283	12	AC079476	329	20.6	64.4	607	2	C0701376	Sequence
C 257	20.8	65.0	218908	12	AC163174	330	20.6	64.4	614	2	C0671277	Sequence
258	20.8	65.0	228645	6	AL450397	331	20.6	64.4	747	5	CR456773	Sequence
259	20.8	65.0	231119	12	AC113521	332	20.6	64.4	747	8	AY892427	Sequence
260	20.8	65.0	234795	6	AC099934	333	20.6	64.4	810	2	AR380730	Sequence
261	20.8	65.0	237247	6	AC095563	334	20.6	64.4	810	5	HUMRPL7X	Sequence
262	20.8	65.0	243689	12	AC095265	335	20.6	64.4	818	2	C0721654	Sequence
C 263	20.8	65.0	248935	12	AC135200	336	20.6	64.4	825	5	HSRPL7	Sequence
C 264	20.8	65.0	249083	12	AC098753	337	20.6	64.4	827	5	HSRBRPL7A	Sequence
C 265	20.8	65.0	258038	12	AC016018	338	20.6	64.4	830	5	BC071671	Sequence
C 266	20.8	65.0	258226	12	AC167453	339	20.6	64.4	831	5	BC087837	Sequence
267	20.8	65.0	259263	12	AC165522	340	20.6	64.4	833	5	BC071895	Sequence
C 268	20.8	65.0	260162	12	AC127763	341	20.6	64.4	838	2	CS162870	Sequence
C 269	20.8	65.0	318115	12	AC098550	342	20.6	64.4	838	5	BC006095	Sequence
C 270	20.6	64.4	102	2	AX322263	343	20.6	64.4	839	5	BC071894	Sequence
C 271	20.6	64.4	126	2	AX523086	344	20.6	64.4	861	5	HSRBRPL7B	Sequence
272	20.6	64.4	105	2	C0734682	345	20.6	64.4	866	5	BC009599	Sequence
273	20.6	64.4	205	5	BC070215	346	20.6	64.4	882	2	C0481861	Sequence
274	20.6	64.4	211	2	C0701109	347	20.6	64.4	892	5	BC008850	Sequence
275	20.6	64.4	234	2	AC062414	348	20.6	64.4	1365	2	AR448326	Sequence
276	20.6	64.4	265	2	AX985386	349	20.6	64.4	1383	2	BD268056	Sequence
277	20.6	64.4	265	2	BD120245	350	20.6	64.4	1383	2	115237	Sequence
278	20.6	64.4	265	2	AR424692	351	20.6	64.4	1447	5	BC106056	Sequence
279	20.6	64.4	279	2	C0736273	352	20.6	64.4	1841	5	HUMRPL7X	Sequence
280	20.6	64.4	295	2	C0689212	353	20.6	64.4	2219	2	CS071093	Sequence
281	20.6	64.4	316	2	C0680657	354	20.6	64.4	2339	2	AR448325	Sequence
282	20.6	64.4	329	2	C0671400	355	20.6	64.4	23401	5	AC099061	Sequence
283	20.6	64.4	343	2	C0670851	356	20.6	64.4	86829	5	AC011382	Sequence
284	20.6	64.4	356	2	C0667011	357	20.6	64.4	102008	12	AC016552	Sequence
285	20.6	64.4	369	2	C0706324	358	20.6	64.4	107381	12	AC115093	Sequence
286	20.6	64.4	372	2	C0678383	359	20.6	64.4	110000	5	AB128049_17	Sequence
287	20.6	64.4	375	2	C0677902	360	20.6	64.4	110000	15	AB017143_00	Sequence
288	20.6	64.4	383	2	CQ480497	361	20.6	64.4	115932	5	HS1107	Sequence
289	20.6	64.4	386	2	C0664882	362	20.6	64.4	121265	12	AF286885	Sequence
290	20.6	64.4	387	2	CQ712454	363	20.6	64.4	127836	5	AC092828	Sequence
291	20.6	64.4	400	2	C0664183	364	20.6	64.4	129109	5	AL3590239	Sequence
292	20.6	64.4	400	2	C0676176	365	20.6	64.4	129592	12	AC170060	Sequence
293	20.6	64.4	406	2	AX360574	366	20.6	64.4	135213	5	AC111149	Sequence
294	20.6	64.4	415	2	C0662299	367	20.6	64.4	13969	5	AL357552	Sequence
295	20.6	64.4	419	2	C0685720	368	20.6	64.4	147098	5	AC008651	Sequence
C 296	20.6	64.4	439	2	CQ472692	369	20.6	64.4	150124	12	AC148886	Sequence
C 297	20.6	64.4	439	2	C0673749	370	20.6	64.4	152747	5	AL359260	Sequence
298	20.6	64.4	440	2	C0707670	371	20.6	64.4	157090	12	AC144497	Sequence
299	20.6	64.4	441	2	C0698538	372	20.6	64.4	161795	12	AC060832	Sequence
300	20.6	64.4	443	2	CQ707833	373	20.6	64.4	167036	12	AC188885	Sequence
301	20.6	64.4	444	2	CQ679348	374	20.6	64.4	170048	12	AC025240	Sequence
C 302	20.6	64.4	452	2	CQ503001	375	20.6	64.4	174346	5	AC108173	Sequence
C 303	20.6	64.4	452	2	CQ51852	376	20.6	64.4	174766	5	AC148659	Sequence
C 304	20.6	64.4	454	2	C0682304	377	20.6	64.4	175150	12	AC136657	Sequence
305	20.6	64.4	462	2	C0692098	378	20.6	64.4	175150	12	AC119777	Sequence
306	20.6	64.4	464	2	C0685059	379	20.6	64.4	176063	12	AC169003	Sequence
307	20.6	64.4	464	2	CQ711448	380	20.6	64.4	181948	12	AC169003	Sequence
308	20.6	64.4	466	2	CQ694492	381	20.6	64.4	186516	6	AC122499	Sequence
309	20.6	64.4	469	2	CQ691352	382	20.6	64.4	191318	12	AC020670	Sequence
310	20.6	64.4	471	2	CQ679552	383	20.6	64.4	197103	5	AL590036	Sequence

C 384	20.6	64.4	203261	12	AC021464	Homo sapi	457	20	62.5	237138	12	AC167679	AC167679 Bos tauru
385	20.6	64.4	222701	12	AC174851	Colobus g	458	20	62.5	237428	12	AC127053	AC127053 Rattus no
386	20.6	64.4	234611	12	AC174629	Colobus g	459	20	62.5	245389	12	AC119463	AC119463 Rattus no
387	20.6	64.4	237624	12	AC113662	Rattus no	460	19.8	61.9	470	2	CO071366	CO071366 Sequence
388	20.6	64.4	244525	5	AC008499	Homo sapi	461	19.8	61.9	566	1	AY696585	AY696585 Pagrus ma
C 389	20.6	64.4	247917	12	AC095906	Rattus no	462	19.8	61.9	756	2	BD216963	BD216963 Novel hum
C 390	20.6	64.4	252132	12	AC095221	Rattus no	463	19.8	61.9	768	7	AR770553	AR770553 Sequence
C 391	20.6	64.4	252630	12	AC112099	Rattus no	464	19.8	61.9	756	2	BV669265	BV669265 S215P6064
C 392	20.6	64.4	256245	12	AC096122	Rattus no	465	19.8	61.9	4065	15	AF628731	AF628731 Rickettsi
C 393	20.6	64.4	255762	12	AC106597	Rattus no	466	19.8	61.9	4734	15	RSP582615	RSP582615 Rickettsi
C 394	20.6	64.4	272658	12	AC131544	Rattus no	467	19.8	61.9	4734	15	RSP582616	RSP582616 Rickettsi
C 395	20.6	64.4	294286	12	AC137391	Rattus no	468	19.8	61.9	4735	15	RSP582613	RSP582613 Rickettsi
C 396	20.6	64.4	297844	12	AC097020	Rattus no	469	19.8	61.9	51049	12	AC155140	AC155140 Bos tauru
C 397	20.6	63.7	348986	15	BX572093	Prochiloto	470	19.8	61.9	53045	6	AL844890	AL844890 Mouse DNA
C 398	20.4	63.7	1648	4	AB072566	Drosoph	471	19.8	61.9	57594	12	AC104552	AC104552 Mus muscu
C 399	20.4	63.7	1708	4	AY096115	Drosoph	472	19.8	61.9	82485	14	AC149769	AC149769 Bos tauru
C 400	20.4	63.7	2595	2	AX653293	Sequence	473	19.8	61.9	84982	5	AY589041	AY589041 Macaca mu
C 401	20.4	63.7	56154	12	AC079961	Homo sapi	474	19.8	61.9	86763	12	AC016406	AC016406 Homo sapi
C 402	20.4	63.7	110000	4	AB017342_03	Continuation (4 of	475	19.8	61.9	110000	15	CR931997_06	Continuation (7 of
C 403	20.4	63.7	110000	4	AB017342_04	Continuation (5 of	476	19.8	61.9	119004	5	AC010230	AC010230 Homo sapi
C 404	20.4	63.7	110000	4	AP008207_328	Continuation (329	477	19.8	61.9	124269	12	AC152482	AC152482 Desypus n
C 405	20.4	63.7	120924	12	AC151492	Continuation (129	478	19.8	61.9	125967	11	AC174721	AC174721 Gasteroste
C 406	20.4	63.7	134736	12	AC151493	Sorex ara	479	19.8	61.9	126474	5	HS181N1	HS181N1 Human DNA
C 407	20.4	63.7	148978	5	HSAP9612	Homo sapi	480	19.8	61.9	136436	12	AC152155	AC152155 Desypus n
C 408	20.4	63.7	149000	5	AC127540	Homo sapi	481	19.8	61.9	147811	11	CR847546	CR847546 Zebrafish
C 409	20.4	63.7	156633	4	AP003344	Oryza sat	482	19.8	61.9	150884	11	BX005161	BX005161 Zebrafish
C 410	20.4	63.7	158939	12	AC102266	Mus muscu	483	19.8	61.9	154727	12	AC152368	AC152368 Desypus n
C 411	20.4	63.7	158950	5	AL391425	Human DNA	484	19.8	61.9	156042	12	AC013614	AC013614 Homo sapi
C 412	20.4	63.7	152276	11	CR388052	Zebrafish	485	19.8	61.9	158311	12	AC009589	AC009589 Homo sapi
C 413	20.4	63.7	153421	12	AC012664	Homo sapi	486	19.8	61.9	158979	6	AL869892	AL869892 Mouse DNA
C 414	20.4	63.7	155777	6	AL807741	Mouse DNA	487	19.8	61.9	159797	5	AC104697	AC104697 Homo sapi
C 415	20.4	63.7	155857	11	BX601643	Zebrafish	488	19.8	61.9	162278	12	AC106761	AC106761 Homo sapi
C 416	20.4	63.7	169931	5	AC005822	Homo sapi	489	19.8	61.9	165345	6	CR932808	CR932808 Mus muscu
C 417	20.4	63.7	176435	12	AC149874	Homo sapi	490	19.8	61.9	165525	6	CR932798	CR932798 Mus muscu
C 418	20.4	63.7	186710	12	AC178132	Strongylo	491	19.8	61.9	178072	6	AC131791	AC131791 Homo sapi
C 419	20.4	63.7	197709	5	AC073641	Homo sapi	492	19.8	61.9	180252	12	AC144538	AC144538 Mouse DNA
C 420	20.4	63.7	199819	12	AC161639	Bos tauru	493	19.8	61.9	181794	6	AL8445186	AL844518 Bos tauru
C 421	20.4	63.7	210885	12	AC012460	Homo sapi	494	19.8	61.9	182023	5	AC055869	AC055869 Homo sapi
C 422	20.4	63.7	211027	12	AC134724	Rattus no	495	19.8	61.9	189246	12	AC177713	AC177713 Strongylo
C 423	20.4	63.7	212486	6	AC153562	Mus muscu	496	19.8	61.9	191954	12	AC159845	AC159845 Bos tauru
C 424	20.4	63.7	213586	12	AC174220	Bos tauru	497	19.8	61.9	192944	12	AC146902	AC146902 Callipebu
C 425	20.4	63.7	220679	6	AL928658	Mouse DNA	498	19.8	61.9	196206	5	AC012368	AC012368 Homo sapi
C 426	20.4	63.7	233339	12	AC094134	Rattus no	499	19.8	61.9	199027	12	AC137061	AC137061 Bos tauru
C 427	20.4	63.7	237003	12	AC172758	Bos tauru	500	19.8	61.9	203840	5	AP0011007	AP0011007 Homo sapi
C 428	20.4	63.7	253322	12	AC110626	Rattus no	501	19.8	61.9	204132	12	AC150925	AC150925 Bos tauru
C 429	20.4	63.7	254812	12	AC107123	Rattus no	502	19.8	61.9	214740	12	AC150925	AC150925 Bos tauru
C 430	20.4	63.7	251576	12	AC127398	Rattus no	503	19.8	61.9	220157	12	AC166843	AC166843 Bos tauru
C 431	20.4	63.7	273967	12	AC160690	Bos tauru	504	19.8	61.9	230158	12	AC159988	AC159988 Bos tauru
C 432	20.4	63.7	273967	12	AC160690	Bos tauru	505	19.8	61.9	235516	12	AC163072	AC163072 Bos tauru
C 433	20.4	63.7	285280	12	AC156819	Bos tauru	506	19.8	61.9	241161	12	AC164734	AC164734 Bos tauru
C 434	20.2	63.1	285280	12	AC156819	Bos tauru	507	19.8	61.9	241237	12	AC156960	AC156960 Bos tauru
C 435	20.2	63.1	96202	5	HSPA26R5	Homo sapi	508	19.8	61.9	242359	6	AC158534	AC158534 Mus muscu
C 436	20.2	63.1	197835	12	AL138824	Human DNA	509	19.8	61.9	245606	12	AC149739	AC149739 Bos tauru
C 437	20.2	62.5	197835	12	AL138824	Human DNA	510	19.8	61.9	245606	12	AC149739	AC149739 Bos tauru
C 438	20.2	62.5	197835	12	AL138824	Human DNA	511	19.8	61.9	245606	12	AC149739	AC149739 Bos tauru
C 439	20.2	62.5	197835	12	AL138824	Human DNA	512	19.8	61.9	245606	12	AC149739	AC149739 Bos tauru
C 440	20.2	62.5	197835	12	AL138824	Human DNA	513	19.8	61.9	245606	12	AC149739	AC149739 Bos tauru
C 441	20.2	62.5	197835	12	AL138824	Human DNA	514	19.8	61.9	245606	12	AC149739	AC149739 Bos tauru
C 442	20.2	62.5	197835	12	AL138824	Human DNA	515	19.8	61.9	245606	12	AC149739	AC149739 Bos tauru
C 443	20.2	62.5	197835	12	AL138824	Human DNA	516	19.8	61.9	245606	12	AC149739	AC149739 Bos tauru
C 444	20.2	62.5	197835	12	AL138824	Human DNA	517	19.8	61.9	245606	12	AC149739	AC149739 Bos tauru
C 445	20.2	62.5	197835	12	AL138824	Human DNA	518	19.8	61.9	245606	12	AC149739	AC149739 Bos tauru
C 446	20.2	62.5	197835	12	AL138824	Human DNA	519	19.8	61.9	245606	12	AC149739	AC149739 Bos tauru
C 447	20.2	62.5	197835	12	AL138824	Human DNA	520	19.8	61.9	245606	12	AC149739	AC149739 Bos tauru
C 448	20.2	62.5	197835	12	AL138824	Human DNA	521	19.8	61.9	245606	12	AC149739	AC149739 Bos tauru
C 449	20.2	62.5	197835	12	AL138824	Human DNA	522	19.8	61.9	245606	12	AC149739	AC149739 Bos tauru
C 450	20.2	62.5	197835	12	AL138824	Human DNA	523	19.8	61.9	245606	12	AC149739	AC149739 Bos tauru
C 451	20.2	62.5	197835	12	AL138824	Human DNA	524	19.8	61.9	245606	12	AC149739	AC149739 Bos tauru
C 452	20.2	62.5	197835	12	AL138824	Human DNA	525	19.8	61.9	245606	12	AC149739	AC149739 Bos tauru
C 453	20.2	62.5	197835	12	AL138824	Human DNA	526	19.8	61.9	245606	12	AC149739	AC149739 Bos tauru
C 454	20.2	62.5	197835	12	AL138824	Human DNA	527	19.8	61.9	245606	12	AC149739	AC149739 Bos tauru
C 455	20.2	62.5	197835	12	AL138824	Human DNA	528	19.8	61.9	245606	12	AC149739	AC149739 Bos tauru
C 456	20.2	62.5	197835	12	AL138824	Human DNA	529	19.8	61.9	245606	12	AC149739	AC149739 Bos tauru

C 530	19.4	60.6	9121	14	OCTRICHA	Z19092 O. cuniculus	C 603	19.4	60.6	193067	12	AC152082	AC152082 Bos tauru
531	19.4	60.6	12350	15	AF522077	AF522077 Lactobacil	604	19.4	60.6	193721	12	AC159081	AC159081 Bos tauru
532	19.4	60.6	12499	12	AC017696	AC017696 Drosophila	605	19.4	60.6	193756	12	AC166397	AC166397 Bos tauru
533	19.4	60.6	16139	15	LPATOVGNS	X94434 L. plantarum	C 606	19.4	60.6	195741	12	AC154156	AC154156 Oryctolag
C 534	19.4	60.6	23574	5	BX005030	BX005030 Human DNA	607	19.4	60.6	196005	12	AC151142	AC151142 Bos tauru
535	19.4	60.6	52423	12	AC100540	AC100540 Mus muscu	C 608	19.4	60.6	196005	12	AC151142	AC151142 Bos tauru
C 536	19.4	60.6	52423	12	AC100540	AC100540 Mus muscu	609	19.4	60.6	196596	6	AC151293	AC151293 Mus muscu
537	19.4	60.6	55506	12	AC032975	AC032975 Homo sapi	C 610	19.4	60.6	197215	12	AC167793	AC167793 Bos tauru
538	19.4	60.6	60583	12	AC036123	AC036123 Homo sapi	C 611	19.4	60.6	197611	12	AC174573	AC174573 Bos tauru
539	19.4	60.6	62866	12	AC175511	AC175511 Bos tauru	612	19.4	60.6	198170	5	AC117394	AC117394 Homo sapi
C 540	19.4	60.6	62866	12	AC175511	AC175511 Bos tauru	613	19.4	60.6	199164	12	AC167225	AC167225 Bos tauru
541	19.4	60.6	65026	12	AC154948	AC154948 Bos tauru	614	19.4	60.6	200574	12	AC171724	AC171724 Bos tauru
C 542	19.4	60.6	72586	6	AL732542	AL732542 Mouse DNA	C 615	19.4	60.6	201748	12	AC161914	AC161914 Bos tauru
C 543	19.4	60.6	82686	12	AC164462	AC164462 Bos tauru	616	19.4	60.6	201820	12	AC166860	AC166860 Oryctolag
544	19.4	60.6	86985	5	BS000638	BS000638 Pan trogl	C 617	19.4	60.6	202307	12	AC164230	AC164230 Bos tauru
C 545	19.4	60.6	87140	12	AC165612	AC165612 Bos tauru	C 618	19.4	60.6	202929	13	AC007882	AC007882 Drosophila
546	19.4	60.6	88294	12	AC162111	AC162111 Bos tauru	619	19.4	60.6	203040	12	AC154952	AC154952 Bos tauru
547	19.4	60.6	93004	5	BS000660	BS000660 Pan trogl	C 620	19.4	60.6	203646	12	AC174472	AC174472 Bos tauru
C 548	19.4	60.6	99901	12	AC169332	AC169332 Bos tauru	621	19.4	60.6	203696	14	AC087860	AC087860 Bos tauru
549	19.4	60.6	105986	12	AC173313	AC173313 Bos tauru	C 622	19.4	60.6	203750	6	AC133516	AC133516 Mus muscu
C 550	19.4	60.6	107558	12	AC025862	AC025862 Homo sapi	C 623	19.4	60.6	205030	6	AL669911	AL669911 Mouse DNA
C 551	19.4	60.6	109757	5	AL162379	AL162379 Human DNA	C 624	19.4	60.6	205594	12	AC169300	AC169300 Bos tauru
C 552	19.4	60.6	110000	15	CP000013_4	Continuation (22 o	C 625	19.4	60.6	206290	12	AC150958	AC150958 Bos tauru
553	19.4	60.6	110000	15	CP000013_05	Continuation (6 of	C 626	19.4	60.6	207408	12	AC167509	AC167509 Bos tauru
C 554	19.4	60.6	110000	15	CP000013_05	Continuation (6 of	627	19.4	60.6	207656	12	AC157055	AC157055 Bos tauru
C 555	19.4	60.6	112129	14	AC162125	AC162125 Oryctolag	628	19.4	60.6	207755	12	AC182014	AC182014 Oryctolag
556	19.4	60.6	122114	12	AC167027	AC167027 Bos tauru	629	19.4	60.6	210361	12	AC152649	AC152649 Bos tauru
C 557	19.4	60.6	136282	12	AC178283	AC178283 Strongylo	C 630	19.4	60.6	212016	12	AC176157	AC176157 Strongylo
558	19.4	60.6	139840	12	AC171415	AC171415 Bos tauru	C 631	19.4	60.6	212461	12	AC165392	AC165392 Oryctolag
C 559	19.4	60.6	140794	12	AC069415	AC069415 Homo sapi	632	19.4	60.6	212928	12	AC172838	AC172838 Bos tauru
560	19.4	60.6	143177	12	AC172997	AC172997 Bos tauru	C 633	19.4	60.6	213252	12	AC153601	AC153601 Bos tauru
561	19.4	60.6	144454	5	BS000637	BS000637 Pan trogl	634	19.4	60.6	214278	12	AC165102	AC165102 Bos tauru
C 562	19.4	60.6	144880	12	AC116434	AC116434 Bos tauru	C 635	19.4	60.6	214762	6	AC167023	AC167023 Mus muscu
C 563	19.4	60.6	145383	12	AC154956	AC154956 Bos tauru	636	19.4	60.6	214785	12	AC153441	AC153441 Bos tauru
C 564	19.4	60.6	150755	6	AC154274	AC154274 Mus muscu	637	19.4	60.6	215472	12	AC173616	AC173616 Bos tauru
565	19.4	60.6	150831	12	AC169331	AC169331 Bos tauru	638	19.4	60.6	216289	12	AC160344	AC160344 Bos tauru
566	19.4	60.6	153845	12	AC173316	AC173316 Strongylo	C 639	19.4	60.6	216485	12	AC165693	AC165693 Bos tauru
C 567	19.4	60.6	155051	6	AC122001	AC122001 Mus muscu	C 640	19.4	60.6	216812	12	AC172223	AC172223 Bos tauru
C 568	19.4	60.6	157016	12	AC174041	AC174041 Bos tauru	C 641	19.4	60.6	219171	12	AC182123	AC182123 Bos tauru
C 569	19.4	60.6	160522	12	AC172253	AC172253 Bos tauru	642	19.4	60.6	220018	12	AC169552	AC169552 Bos tauru
570	19.4	60.6	160551	11	AL954134	AL954134 Zebrafish	643	19.4	60.6	220694	12	AC155178	AC155178 Bos tauru
C 571	19.4	60.6	164049	12	AC151086	AC151086 Bos tauru	C 644	19.4	60.6	221013	12	AC173179	AC173179 Bos tauru
C 572	19.4	60.6	165733	12	AC069072	AC069072 Homo sapi	645	19.4	60.6	221979	12	AC172235	AC172235 Bos tauru
C 573	19.4	60.6	166147	12	AC167080	AC167080 Oryctolag	646	19.4	60.6	222962	12	AC160677	AC160677 Bos tauru
C 574	19.4	60.6	166965	12	AC164495	AC164495 Oryctolag	647	19.4	60.6	223264	12	AC192921	AC192921 Bos tauru
575	19.4	60.6	168416	14	AC149676	AC149676 Bos tauru	C 648	19.4	60.6	224264	13	AE003770	AE003770 Drosophila
C 576	19.4	60.6	168700	12	AC009115	AC009115 Homo sapi	649	19.4	60.6	227186	12	AC151188	AC151188 Bos tauru
577	19.4	60.6	168858	12	AC167507	AC167507 Oryctolag	650	19.4	60.6	227316	12	AC173771	AC173771 Bos tauru
578	19.4	60.6	169242	12	AC175941	AC175941 Strongylo	651	19.4	60.6	227939	12	AC125648	AC125648 Ratine no
C 579	19.4	60.6	170765	5	AL589745	AL589745 Human DNA	C 652	19.4	60.6	228644	12	AC102879	AC102879 Mus muscu
C 580	19.4	60.6	172545	12	AC171287	AC171287 Bos tauru	C 653	19.4	60.6	228869	6	BX470109	BX470109 Mouse DNA
C 581	19.4	60.6	172710	6	AL732584	AL732584 Mouse DNA	C 654	19.4	60.6	228945	12	AC170306	AC170306 Bos tauru
C 582	19.4	60.6	172891	12	AC084011	AC084011 Homo sapi	C 655	19.4	60.6	228970	12	AC172682	AC172682 Bos tauru
C 583	19.4	60.6	174420	5	AC010723	AC010723 Homo sapi	656	19.4	60.6	229218	12	AC107523	AC107523 Ratine no
584	19.4	60.6	174996	14	AC150881	AC150881 Bos tauru	657	19.4	60.6	229318	12	AC159862	AC159862 Bos tauru
C 585	19.4	60.6	175124	13	AC007885	AC007885 Drosophila	658	19.4	60.6	230413	12	AC169583	AC169583 Bos tauru
C 586	19.4	60.6	178791	13	AC008218	AC008218 Drosophila	C 659	19.4	60.6	231452	12	AC160482	AC160482 Bos tauru
587	19.4	60.6	179821	5	AC147144	AC147144 Pan trogl	660	19.4	60.6	233492	12	AC111723	AC111723 Ratine no
C 588	19.4	60.6	179821	5	AC147144	AC147144 Pan trogl	C 661	19.4	60.6	233716	12	AC163302	AC163302 Bos tauru
C 589	19.4	60.6	179996	12	AC163503	AC163503 Bos tauru	662	19.4	60.6	236554	12	AC156895	AC156895 Bos tauru
C 590	19.4	60.6	180245	12	AC147729	AC147729 Oryctolag	C 663	19.4	60.6	237096	12	AC131974	AC131974 Mus muscu
C 591	19.4	60.6	180952	12	AC163228	AC163228 Bos tauru	664	19.4	60.6	237552	12	AC094472	AC094472 Ratine no
C 592	19.4	60.6	182892	12	AC180268	AC180268 Strongylo	665	19.4	60.6	237752	12	AC155733	AC155733 Bos tauru
C 593	19.4	60.6	183681	12	AC170690	AC170690 Bos tauru	C 666	19.4	60.6	239566	5	HSABJ0147	HSABJ0147 Homo sapi
C 594	19.4	60.6	184774	12	AC170476	AC170476 Bos tauru	667	19.4	60.6	239705	12	AC163789	AC163789 Bos tauru
C 595	19.4	60.6	185345	12	AC163826	AC163826 Bos tauru	C 668	19.4	60.6	239705	12	AC163789	AC163789 Bos tauru
596	19.4	60.6	185881	12	AC148860	AC148860 Oryctolag	C 669	19.4	60.6	239750	12	AC165571	AC165571 Bos tauru
597	19.4	60.6	186081	12	AC167821	AC167821 Bos tauru	670	19.4	60.6	240970	12	AC177544	AC177544 Bos tauru
598	19.4	60.6	187206	12	AC152436	AC152436 Bos tauru	C 671	19.4	60.6	241920	12	AC171834	AC171834 Bos tauru
599	19.4	60.6	189026	12	AC159727	AC159727 Bos tauru	C 672	19.4	60.6	242059	12	AC174789	AC174789 Bos tauru
600	19.4	60.6	189347	12	AC162581	AC162581 Bos tauru	C 673	19.4	60.6	243285	12	AC170250	AC170250 Bos tauru
601	19.4	60.6	190086	12	AC172086	AC172086 Bos tauru	C 674	19.4	60.6	245606	12	AC197339	AC197339 Bos tauru
602	19.4	60.6	190650	12	AC157615	AC157615 Oryctolag	675	19.4	60.6	247489	12	AC162999	AC162999 Bos tauru

676	19.4	60.6	248002	12	AC159043	AC159043 Bos tauru	c 749	19.2	60.0	103125	4	AC005560	AC005560 Arabidops
677	19.4	60.6	248471	12	AC170276	AC170276 Bos tauru	c 750	19.2	60.0	107641	5	AC078921	AC078921 Homo sapi
678	19.4	60.6	249984	12	AC163166	AC163166 Bos tauru	c 751	19.2	60.0	109359	5	AC004857	AC004857 Homo sapi
c 679	19.4	60.6	250150	12	AC168109	AC168109 Bos tauru	c 752	19.2	60.0	110000	4	CR382128_03	Continuation (4 of
680	19.4	60.6	250589	12	AC130139	AC130139 Rattus no	c 753	19.2	60.0	110000	4	CR382128_04	Continuation (5 of
681	19.4	60.6	252830	12	AC170906	AC170906 Mus muscu	c 754	19.2	60.0	110000	2	AC138525_3	Continuation (4 of
682	19.4	60.6	263330	12	AC156186	AC156186 Bos tauru	c 755	19.2	60.0	110000	12	AC073744_3	Continuation (4 of
683	19.4	60.6	263390	12	AC172070	AC172070 Bos tauru	c 756	19.2	60.0	110000	13	AE003597_0	AB003597 Drosophi1
c 684	19.4	60.6	263390	12	AC172070	AC172070 Bos tauru	c 757	19.2	60.0	110000	15	CP000252_04	Continuation (5 of
c 685	19.4	60.6	264142	12	AC163812	AC163812 Bos tauru	c 758	19.2	60.0	110000	15	AE016795_30	Continuation (31 o
686	19.4	60.6	264394	12	AC169574	AC169574 Bos tauru	c 759	19.2	60.0	111073	5	AC006375	AC006375 Homo sapi
687	19.4	60.6	265305	12	AC172071	AC172071 Bos tauru	c 760	19.2	60.0	112786	11	BX10370	BX10370 Zebrafish
c 688	19.4	60.6	265654	12	AC161963	AC161963 Bos tauru	c 761	19.2	60.0	116099	12	AC168919	AC168919 Homo sapi
c 689	19.4	60.6	266841	12	AC163815	AC163815 Bos tauru	c 762	19.2	60.0	123105	5	AC181950	AC181950 Strongylo
c 690	19.4	60.6	269397	12	AC166961	AC166961 Bos tauru	c 763	19.2	60.0	124635	12	AP000593	AP000593 Homo sapi
691	19.4	60.6	271938	12	AC173391	AC173391 Bos tauru	c 764	19.2	60.0	127735	5	HS257824	AL034424 Human DNA
c 692	19.4	60.6	273527	12	AC156740	AC156740 Bos tauru	c 765	19.2	60.0	129593	10	AF512031	AF512031 Choriastion
693	19.4	60.6	276201	12	AC162184	AC162184 Bos tauru	c 766	19.2	60.0	132700	12	AC040939	AC040939 Homo sapi
c 694	19.4	60.6	282611	12	AC156719	AC156719 Bos tauru	c 767	19.2	60.0	135680	12	AC125299	AC125299 Rattus no
c 695	19.4	60.6	285534	12	AC164124	AC164124 Bos tauru	c 768	19.2	60.0	142911	12	AC178394	AC178394 Strongylo
696	19.4	60.6	288742	12	AC162819	AC162819 Bos tauru	c 769	19.2	60.0	146210	12	BX927333	BX927333 Danto rex
c 697	19.4	60.6	294000	12	AC163910	AC163910 Bos tauru	c 770	19.2	60.0	147140	12	AC136623	AC136623 Homo sapi
c 698	19.4	60.6	295778	12	AC162404	AC162404 Bos tauru	c 771	19.2	60.0	148184	12	AC017777	AC017777 Drosophi1
c 699	19.4	60.6	305362	12	AC159698	AC159698 Bos tauru	c 772	19.2	60.0	149051	12	AC138963	AC138963 Homo sapi
c 700	19.4	60.6	307956	12	AC103312	AC103312 Rattus no	c 773	19.2	60.0	150088	5	AL359925	AL359925 Human DNA
c 701	19.4	60.6	309289	12	AC167679	AC167679 Bos tauru	c 774	19.2	60.0	150879	12	AC134277	AC134277 Rattus no
c 702	19.4	60.6	310420	12	AC174968	AC174968 Bos tauru	c 775	19.2	60.0	152680	12	AC173490	AC173490 Bos tauru
c 703	19.4	60.6	310420	12	AC174968	AC174968 Bos tauru	c 776	19.2	60.0	154677	12	AC016196	AC016196 Homo sapi
704	19.4	60.6	324050	15	AL935253	AL935253 Lactobact	c 777	19.2	60.0	154688	12	AC179610	AC179610 Strongylo
705	19.4	60.6	332904	12	AC116227	AC116227 Rattus no	c 778	19.2	60.0	155335	12	AC005734	AC005734 Drosophi1
c 706	19.4	60.6	333532	12	AC174032	AC174032 Bos tauru	c 779	19.2	60.0	157531	5	AL133409	AL133409 Human DNA
c 707	19.4	60.6	338656	12	AC150935	AC150935 Bos tauru	c 780	19.2	60.0	158324	5	AL136611	AL136611 Homo sapi
708	19.4	60.6	345829	15	BX571868	BX571868 Photorhab	c 781	19.2	60.0	158828	12	AC158895	AC158895 Bos tauru
c 709	19.4	60.6	349980	2	AX770905	AX770905 Sequence	c 782	19.2	60.0	161446	13	AC092233	AC092233 Drosophi1
710	19.2	60.0	228	2	AR449932	AR449932 Sequence	c 783	19.2	60.0	161647	5	AC079080	AC079080 Homo sapi
711	19.2	60.0	450	10	DQ057305	DQ057305 Human pap	c 784	19.2	60.0	162462	12	CR293517	CR293517 Danto rex
712	19.2	60.0	450	10	DQ057308	DQ057308 Human pap	c 785	19.2	60.0	162466	12	AC015643	AC015643 Homo sapi
c 713	19.2	60.0	699	2	MX260058	MX260058 Sequence	c 786	19.2	60.0	162589	12	AL365358	AL365358 Homo sapi
c 714	19.2	60.0	1008	6	MUSJ112PA	MUSJ112PA	c 787	19.2	60.0	163794	12	AC138922	AC138922 Homo sapi
c 715	19.2	60.0	1119	4	AY819048	AY819048	c 788	19.2	60.0	164125	12	AC155155	AC155155 Bos tauru
c 716	19.2	60.0	1157	13	RBH417022	RBH417022	c 789	19.2	60.0	166668	12	AC025258	AC025258 Homo sapi
c 717	19.2	60.0	1271	2	122653	122653 Sequence 14	c 790	19.2	60.0	167285	5	AC129916	AC129916 Homo sapi
c 718	19.2	60.0	1271	2	122653	122653 Sequence 14	c 791	19.2	60.0	167315	12	AC137501	AC137501 Homo sapi
c 719	19.2	60.0	1776	2	AR449906	AR449906 Sequence	c 792	19.2	60.0	168862	5	AC090532	AC090532 Homo sapi
720	19.2	60.0	2592	11	BC067327	BC067327 Xenopus t	c 793	19.2	60.0	169660	12	AC147854	AC147854 Homo sapi
721	19.2	60.0	2680	6	BC023140	BC023140 Mus muscu	c 794	19.2	60.0	169760	12	AC073619	AC073619 Homo sapi
c 722	19.2	60.0	3721	14	BFOCT1	BF022986 Bos tauru	c 795	19.2	60.0	170901	5	AC018878	AC018878 Homo sapi
723	19.2	60.0	4406	2	CQ572304	CQ572304 Sequence	c 796	19.2	60.0	171285	12	AC158071	AC158071 Bos tauru
724	19.2	60.0	5891	12	AC020186	AC020186 Drosophi1	c 797	19.2	60.0	173716	5	AC115090	AC115090 Homo sapi
725	19.2	60.0	5893	13	AY094730	AY094730 Drosophi1	c 798	19.2	60.0	174311	12	AL165319	AL165319 Homo sapi
726	19.2	60.0	5986	6	AK129055	AK129055 Mus muscu	c 799	19.2	60.0	174340	12	AC165306	AC165306 Bos tauru
727	19.2	60.0	7546	2	CQ572303	CQ572303 Sequence	c 800	19.2	60.0	177163	5	AC138904	AC138904 Homo sapi
728	19.2	60.0	7700	10	AB027020	AB027020 Human pap	c 801	19.2	60.0	177203	12	AC024531	AC024531 Homo sapi
c 729	19.2	60.0	10719	15	AE009067	AE009067 Agrobacte	c 802	19.2	60.0	178127	12	AC079976	AC079976 Homo sapi
c 730	19.2	60.0	10781	15	AE008033	AE008033 Agrobacte	c 803	19.2	60.0	178213	12	AC157386	AC157386 Bos tauru
731	19.2	60.0	29159	12	AC137965	AC137965 Homo sapi	c 804	19.2	60.0	180366	6	AC162448	AC162448 Mus muscu
732	19.2	60.0	30303	12	AC014888	AC014888 Drosophi1	c 805	19.2	60.0	180588	11	AL1954672	AL1954672 Zebrafish
733	19.2	60.0	37909	5	AP005232	AP005232 Homo sapi	c 806	19.2	60.0	181217	12	AC169747	AC169747 Bos tauru
734	19.2	60.0	40327	5	AP005379	AP005379 Homo sapi	c 807	19.2	60.0	184461	12	AC173062	AC173062 Bos tauru
735	19.2	60.0	43916	6	CR753826	CR753826 Mouse DNA	c 808	19.2	60.0	184893	12	AC153164	AC153164 Bos tauru
c 736	19.2	60.0	44260	12	AC157742	AC157742 Xenopus t	c 809	19.2	60.0	185165	12	AC134309	AC134309 Homo sapi
c 737	19.2	60.0	56531	12	AC105232	AC105232 Homo sapi	c 810	19.2	60.0	185237	13	AC172095	AC172095 Bos tauru
c 738	19.2	60.0	69349	12	AC175420	AC175420 Bos tauru	c 811	19.2	60.0	188359	13	AC008189	AC008189 Drosophi1
c 739	19.2	60.0	73774	12	CR407561_3	CR407561_3	c 812	19.2	60.0	188491	12	AC170548	AC170548 Bos tauru
c 740	19.2	60.0	89047	2	AR408755	AR408755 Sequence	c 813	19.2	60.0	188994	12	AC150955	AC150955 Bos tauru
c 741	19.2	60.0	89047	2	AX067459	AX067459 Sequence	c 814	19.2	60.0	189161	5	AL339715	AL339715 Human DNA
742	19.2	60.0	90458	5	AL355552	AL355552 Human DNA	c 815	19.2	60.0	189326	6	AL1607030	AL1607030 Mouse DNA
743	19.2	60.0	90448	12	AC167829	AC167829 Bos tauru	c 816	19.2	60.0	189351	12	AC156120	AC156120 Bos tauru
c 744	19.2	60.0	91868	12	AC172665	AC172665 Bos tauru	c 817	19.2	60.0	189441	5	AC182176	AC182176 Bos tauru
c 745	19.2	60.0	96255	12	AC153676	AC153676 Bos tauru	c 818	19.2	60.0	192841	12	AC007597	AC007597 Homo sapi
c 746	19.2	60.0	98660	12	AC153664	AC153664 Bos tauru	c 819	19.2	60.0	193772	12	AC016833	AC016833 Homo sapi
c 747	19.2	60.0	101006	12	AC079392	AC079392 Homo sapi	c 820	19.2	60.0	195757	13	AC010009	AC010009 Drosophi1
c 748	19.2	60.0	102146	12	AC133396	AC133396 Felis cat	c 821	19.2	60.0	195909	12	AC150457	AC150457 Callithi1

C 822	19.2	60.0	197658	12	AC124208	AC124208 Rattus no	895	19.2	60.0	284404	12	AC160044	AC160044 Bos tauru
C 823	19.2	60.0	197679	6	AC128663	AC128663 Mus muscu	896	19.2	60.0	287164	12	AC169365	AC169365 Bos tauru
C 824	19.2	60.0	197913	6	AC173177	AC173177 Bos tauru	C 897	19.2	60.0	287927	12	AC079530	AC079530 Mus muscu
C 825	19.2	60.0	198033	12	AC123973	AC123973 Canis fam	C 898	19.2	60.0	226111	12	AC120083	AC120083 Rattus no
C 826	19.2	60.0	199647	12	AC150632	AC150632 Bos tauru	C 899	19.2	60.0	298908	12	AC167569	AC167569 Bos tauru
C 827	19.2	60.0	199785	5	AL355512	AL355512 Human DNA	C 900	19.2	60.0	301540	12	AC156687	AC156687 Bos tauru
C 828	19.2	60.0	199921	12	AC153187	AC153187 Bos tauru	C 901	19.2	60.0	314150	15	CJ11168X1	AL139074 Campyloba
C 829	19.2	60.0	200167	12	AC152658	AC152658 Bos tauru	C 902	19.2	60.0	328861	12	AC162688	AC162688 Bos tauru
C 830	19.2	60.0	201524	6	AC161246	AC161246 Mus muscu	C 903	19.2	60.0	329032	12	AC159851	AC159851 Bos tauru
C 831	19.2	60.0	201582	12	AC127484	AC127484 Homo sapi	C 904	19.2	60.0	329556	12	AC128866	AC128866 Rattus no
C 832	19.2	60.0	204444	12	AL445436	AL445436 Homo sapi	C 905	19.2	59.4	204	2	CQ73668	CQ73668 Sequence
C 833	19.2	60.0	205733	12	AC171300	AC171300 Bos tauru	C 906	19.2	59.4	294	2	CQ655489	CQ655489 Sequence
C 834	19.2	60.0	206433	5	AC147978	AC147978 Pan trogl	C 907	19.2	59.4	348	2	CQ080781	CQ080781 Sequence
C 835	19.2	60.0	207156	11	CR405694	CR405694 Zebrafish	C 908	19.2	59.4	348	2	CQ114449	CQ114449 Sequence
C 836	19.2	60.0	207164	12	AC171484	AC171484 Bos tauru	C 909	19.2	59.4	348	2	CQ115244	CQ115244 Sequence
C 837	19.2	60.0	207164	12	AC171484	AC171484 Bos tauru	C 910	19.2	59.4	348	2	CQ153328	CQ153328 Sequence
C 838	19.2	60.0	208233	6	AC114425	AC114425 Mus muscu	C 911	19.2	59.4	348	2	CQ154075	CQ154075 Sequence
C 839	19.2	60.0	208357	12	AC165662	AC165662 Bos tauru	C 912	19.2	59.4	348	2	CQ186246	CQ186246 Sequence
C 840	19.2	60.0	208793	12	AC164393	AC164393 Bos tauru	C 913	19.2	59.4	348	2	CQ186656	CQ186656 Sequence
C 841	19.2	60.0	209551	12	AC174533	AC174533 Bos tauru	C 914	19.2	59.4	348	2	CQ236617	CQ236617 Sequence
C 842	19.2	60.0	210209	12	AC164216	AC164216 Bos tauru	C 915	19.2	59.4	348	2	CQ237310	CQ237310 Sequence
C 843	19.2	60.0	210591	12	AC166908	AC166908 Bos tauru	C 916	19.2	59.4	348	2	CQ274214	CQ274214 Sequence
C 844	19.2	60.0	211526	12	AC164059	AC164059 Bos tauru	C 917	19.2	59.4	348	2	CQ274934	CQ274934 Sequence
C 845	19.2	60.0	212118	12	AC157062	AC157062 Bos tauru	C 918	19.2	59.4	348	2	CQ311384	CQ311384 Sequence
C 846	19.2	60.0	212852	12	AC167799	AC167799 Bos tauru	C 919	19.2	59.4	348	2	CQ311908	CQ311908 Sequence
C 847	19.2	60.0	213515	12	AC145231	AC145231 Homo sapi	C 920	19.2	59.4	348	2	CQ348522	CQ348522 Sequence
C 848	19.2	60.0	214786	5	AC007610	AC007610 Homo sapi	C 921	19.2	59.4	348	2	CQ349281	CQ349281 Sequence
C 849	19.2	60.0	214966	12	AC163205	AC163205 Bos tauru	C 922	19.2	59.4	348	2	CQ734687	CQ734687 Sequence
C 850	19.2	60.0	216035	12	AC152780	AC152780 Bos tauru	C 923	19.2	59.4	454	2	CQ712518	CQ712518 Sequence
C 851	19.2	60.0	217992	12	AC164897	AC164897 Bos tauru	C 924	19.2	59.4	470	2	CQ685505	CQ685505 Sequence
C 852	19.2	60.0	218265	12	AC129866	AC129866 Bos tauru	C 925	19.2	59.4	527	2	CQ071565	CQ071565 Sequence
C 853	19.2	60.0	218394	12	AC153179	AC153179 Bos tauru	C 926	19.2	59.4	527	2	CQ101414	CQ101414 Sequence
C 854	19.2	60.0	218749	12	AC170984	AC170984 Bos tauru	C 927	19.2	59.4	527	2	CQ102165	CQ102165 Sequence
C 855	19.2	60.0	220869	12	AC162215	AC162215 Bos tauru	C 928	19.2	59.4	527	2	CQ140406	CQ140406 Sequence
C 856	19.2	60.0	221291	12	AC173161	AC173161 Bos tauru	C 929	19.2	59.4	527	2	CQ141115	CQ141115 Sequence
C 857	19.2	60.0	221701	12	AC152603	AC152603 Bos tauru	C 930	19.2	59.4	527	2	CQ176396	CQ176396 Sequence
C 858	19.2	60.0	222034	12	AC157219	AC157219 Bos tauru	C 931	19.2	59.4	527	2	CQ176768	CQ176768 Sequence
C 859	19.2	60.0	225629	12	AC169222	AC169222 Bos tauru	C 932	19.2	59.4	527	2	CQ223731	CQ223731 Sequence
C 860	19.2	60.0	225673	12	AC157181	AC157181 Bos tauru	C 933	19.2	59.4	527	2	CQ224384	CQ224384 Sequence
C 861	19.2	60.0	226085	12	AC107125	AC107125 Rattus no	C 934	19.2	59.4	527	2	CQ261719	CQ261719 Sequence
C 862	19.2	60.0	226908	5	AC162060	AC162060 Bos tauru	C 935	19.2	59.4	527	2	CQ262402	CQ262402 Sequence
C 863	19.2	60.0	230510	5	AC084794	AC084794 Homo sapi	C 936	19.2	59.4	527	2	CQ298983	CQ298983 Sequence
C 864	19.2	60.0	230939	12	AC153225	AC153225 Bos tauru	C 937	19.2	59.4	527	2	CQ299461	CQ299461 Sequence
C 865	19.2	60.0	233947	12	AC174810	AC174810 Bos tauru	C 938	19.2	59.4	527	2	CQ335874	CQ335874 Sequence
C 866	19.2	60.0	236066	12	AC163590	AC163590 Bos tauru	C 939	19.2	59.4	527	2	CQ336590	CQ336590 Sequence
C 867	19.2	60.0	236198	12	AC108547	AC108547 Rattus no	C 940	19.2	59.4	535	2	CS246372	CS246372 Sequence
C 868	19.2	60.0	236730	12	AC152745	AC152745 Bos tauru	C 941	19.2	59.4	827	2	CQ724928	CQ724928 Sequence
C 869	19.2	60.0	239570	12	AC137983	AC137983 Mus muscu	C 942	19.2	59.4	1105	6	BC049631	BC049631 Mus muscu
C 870	19.2	60.0	240737	12	AC163429	AC163429 Bos tauru	C 943	19.2	59.4	2235	2	SPC33132	SPC33132 Sequence
C 871	19.2	60.0	241634	12	AC162056	AC162056 Bos tauru	C 944	19.2	59.4	2235	2	AR594990	AR594990 Sequence
C 872	19.2	60.0	242670	12	AC151187	AC151187 Bos tauru	C 945	19.2	59.4	2235	2	AB060586	AB060586 Sequence
C 873	19.2	60.0	245317	12	AC116210	AC116210 Rattus no	C 946	19.2	59.4	2306	4	AB060586	AB060586 Sequence
C 874	19.2	60.0	246770	12	AC096461	AC096461 Rattus no	C 947	19.2	59.4	2369	2	E30091	E30091 Endo-beta-N
C 875	19.2	60.0	249674	12	AC133660	AC133660 Rattus no	C 948	19.2	59.4	2369	2	AR594989	AR594989 Sequence
C 876	19.2	60.0	250029	5	AB014310	AB014310 Homo sapi	C 949	19.2	59.4	3507	2	CQ0851073	CQ0851073 Sequence
C 877	19.2	60.0	250860	12	AC167897	AC167897 Bos tauru	C 950	19.2	59.4	3507	5	AK128281	AK128281 Homo sapi
C 878	19.2	60.0	250899	12	AC167352	AC167352 Bos tauru	C 951	19.2	59.4	20607	15	AB008719	AB008719 Sequence
C 879	19.2	60.0	251166	12	AC163060	AC163060 Bos tauru	C 952	19.2	59.4	39596	12	AL445257	AL445257 Homo sapi
C 880	19.2	60.0	255567	12	AC153455	AC153455 Bos tauru	C 953	19.2	59.4	70648	12	AC174548	AC174548 Strongylo
C 881	19.2	60.0	256662	12	AC174206	AC174206 Bos tauru	C 954	19.2	59.4	78163	12	AC179555	AC179555 Strongylo
C 882	19.2	60.0	257581	12	AC157013	AC157013 Bos tauru	C 955	19.2	59.4	85660	5	AL357148	AL357148 Human DNA
C 883	19.2	60.0	258016	12	AC163323	AC163323 Bos tauru	C 956	19.2	59.4	94984	5	AC114804	AC114804 Homo sapi
C 884	19.2	60.0	261043	13	AB003627	AB003627 Drosophill	C 957	19.2	59.4	95903	15	CP000100_26	CP000100_26 Sequence
C 885	19.2	60.0	264225	13	AB003627	AB003627 Drosophill	C 958	19.2	59.4	99783	12	AC139911	AC139911 Rattus no
C 886	19.2	60.0	268294	12	AC020885	AC020885 Mus muscu	C 959	19.2	59.4	104081	5	HSJD26337	HSJD26337 Human DNA
C 887	19.2	60.0	268949	12	AC164751	AC164751 Bos tauru	C 960	19.2	59.4	104712	5	AL390786	AL390786 Human DNA
C 888	19.2	60.0	269471	12	AC152254	AC152254 Bos tauru	C 961	19.2	59.4	106988	5	AC0001226	AC0001226 Genomic s
C 889	19.2	60.0	270456	12	AC115275	AC115275 Rattus no	C 962	19.2	59.4	109201	5	AC000403	AC000403 Genomic s
C 890	19.2	60.0	271586	12	AC126739	AC126739 Rattus no	C 963	19.2	59.4	110000	12	AC115960_1	AC115960_1 Sequence
C 891	19.2	60.0	272657	12	AC155919	AC155919 Bos tauru	C 964	19.2	59.4	110000	15	AB014613_24	AB014613_24 Sequence
C 892	19.2	60.0	273418	13	AB003605	AB003605 Drosophill	C 965	19.2	59.4	110000	15	AB015928_39	AB015928_39 Sequence
C 893	19.2	60.0	275140	12	AC158015	AC158015 Bos tauru	C 966	19.2	59.4	110000	15	AB017220_06	AB017220_06 Sequence
C 894	19.2	60.0	277757	12	AC159481	AC159481 Bos tauru	C 967	19.2	59.4	110000	15	AP008231_16	AP008231_16 Sequence

C 968	19	59.4	110000	15	BA000004	40	Continuation (41 o
C 969	19	59.4	110000	15	BA000039	24	Continuation (25 o
C 970	19	59.4	110000	15	CP000026	72	Continuation (23 o
C 971	19	59.4	110000	15	CP000026	73	Continuation (24 o
C 972	19	59.4	119956	12	AC154939		Continuation (24 o
C 973	19	59.4	121963	5	AL162337		Continuation (24 o
C 974	19	59.4	132262	12	AL112364		Continuation (24 o
C 975	19	59.4	139687	5	AL590993		Continuation (24 o
C 976	19	59.4	143717	12	AC069245		Continuation (24 o
C 977	19	59.4	143969	5	AL357552		Continuation (24 o
C 978	19	59.4	147007	12	AC157471		Continuation (24 o
C 979	19	59.4	146654	5	AL683813		Continuation (24 o
C 980	19	59.4	153650	5	AC022296		Continuation (24 o
C 981	19	59.4	153927	12	AC155985		Continuation (24 o
C 982	19	59.4	153953	6	AC153577		Continuation (24 o
C 983	19	59.4	154258	12	AC079049		Continuation (24 o
C 984	19	59.4	159862	12	AC160582		Continuation (24 o
C 985	19	59.4	160064	12	AC027235		Continuation (24 o
C 986	19	59.4	160314	12	AC090661		Continuation (24 o
C 987	19	59.4	163475	5	AC078815		Continuation (24 o
C 988	19	59.4	165206	12	AC068488		Continuation (24 o
C 989	19	59.4	165237	2	AX235055		Continuation (24 o
C 990	19	59.4	166305	12	AC160585		Continuation (24 o
C 991	19	59.4	166866	11	BX005238		Continuation (24 o
C 992	19	59.4	166980	12	AC052907		Continuation (24 o
C 993	19	59.4	167865	5	AC087446		Continuation (24 o
C 994	19	59.4	168344	6	AC126028		Continuation (24 o
C 995	19	59.4	169362	12	AL136440		Continuation (24 o
C 996	19	59.4	169669	5	AL157777		Continuation (24 o
C 997	19	59.4	171135	5	AC093895		Continuation (24 o
C 998	19	59.4	172731	12	AC020623		Continuation (24 o
C 999	19	59.4	175831	5	AC124685		Continuation (24 o
C 1000	19	59.4	175064	5	AC074277		Continuation (24 o

ALIGNMENTS

ACCESSION	AR349182
VERSION	AR349182.1
KEYWORDS	GI:3749887
SOURCE	Unknown.
ORGANISM	Unknown.
REFERENCE	Unclassified.
AUTHORS	1 (bases 1 to 32)
TITLE	Carter,N.M.
	Nucleic acid probes complementary to human papilloma virus nucleic acid
JOURNAL	Patent: US 6583278-A 122 24-UTN-2003;
	Gen-Probe Incorporated; San Diego, CA
FEATURES	Location/Qualifiers
source	1..32 /organism="unknown" /mol_type="genomic DNA"
ORIGIN	
Query Match	100.0%; Score 32; DB 2; Length 32;
Best Local Similarity	100.0%; Pred. No. 0.004;
Matches	32; Conservative 0; Mismatches 0; Indels 0; Gaps 0.
CY	1 TTATTAAATGAAGTGCCTGCGGTGCCAAGAACC 32
Db	32 TTATTAAATGAAGTGCCTGCGGTGCCAAGAACC 1
RESULT 3	
LOCUS	AR349183 32 bp. DNA
DEFINITION	Sequence 123 from patent US 6583278.
ACCESSION	AR349183
VERSION	AR349183.1 GI:3749888
	PAT 17-AUG-2003

RESULT 1					
LOCUS	AR349181				
DEFINITION	AR349181 Sequence 121 from patent US 6583278.		32 bp	DNA	linear PAT 17-AUG-2003
ACCESSION	AR349181				
VERSION	AR349181.1		GI:33749886		
KEYWORDS	.				
SOURCE	unknown.				
ORGANISM	unclassified.				
REFERENCE	1 (bases 1 to 32)				
AUTHORS	Carter,N.M,				
TITLE	Nucleic acid probes complementary to human papilloma virus nucleic acid				
JOURNAL	Patent: US 6583278-A 121 24-JUN-2003;				
FEATURES	Gen-Probe Incorporated; San Diego, CA				
SOURCE	Location/Qualifiers				
	1..32				
	/organism="unknown"				
	/mol_type="genomic DNA"				
ORIGIN					
Query Match		100.0%; Score 32; DB 2;	Length 32;		
Best local similarity		100.0%; Pred. No. 0.004;			
Matches 32;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;	
OY	1 TTATTATTAAGGTGCTGGCGGTGCCAATAACC 32 				
Db	1 TTATTATTAAGGTGCTGGCGGTGCCAATAACC 32 				
RESULT 2					
LOCUS	AR349182/c				
DEFINITION	Sequence 122 from patent US 6583278.		32 bp	DNA	linear PAT 17-AUG-2003

```

AUTHORS      CARTER,N.M.
TITLE        Nucleic acid probes complementary to human papilloma virus nucleic
              acid
JOURNAL      Patent: US 6583278-A 123 24-UTN-2003;
              Gen-Probe Incorporated; San Diego, CA
FEATURES     Location/Qualifiers
              1..32
              /organism="unknown"
              /mol_type="genomic DNA"

ORIGIN
Query Match      100.0%; Score 32; DB 2; Length 32;
Best Local Similarity 100.0%; Pident. No. 0.004;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY
1 TTATTATATGAGTGCCTGCGGTGCAGAAACC 32
  |||
1 TTATTATATGAGTGCCTGCGGTGCAGAAACC 32

Db
1 TTATTATATGAGTGCCTGCGGTGCAGAAACC 32

RESULT 4
AR349184/c      32 bp   DNA   linear   PAT 17-AUG-2003
LOCUS           AR349184
DEFINITION      Sequence 124 from patent US 6583278.
ACCESSION       AR349184
VERSION         AR349184.1   GI:33749869
KEYWORDS
SOURCE          Unknown.
ORGANISM        Unknown.
REFERENCE       Unclassified.
AUTHORS         1 (bases 1 to 32)
                Carter,N.M.
TITLE           Nucleic acid probes complementary to human papilloma virus nucleic
                acid
JOURNAL         Patent: US 6583278-A 124 24-UTN-2003;

```

Gen-Probe Incorporated, San Diego, CA

FEATURES
source
1. .32
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 100.0%; Score 32; DB 2; Length 32;
Best Local Similarity 100.0%; Pred. No. 0.004;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TTATTATAAGTGCTCGCGTGCAGAAACC 32
32 TTATTATAAGTGCTCGCGTGCAGAAACC 1

Db 32 TTATTATAAGTGCTCGCGTGCAGAAACC 1

RESULT 5
AR080321 476 bp DNA linear PAT 31-AUG-2000
LOCUS Sequence 16 from patent US 5968761.
DEFINITION AR080321
ACCESSION AR080321 GI:10007056
VERSION AR080321.1
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 476)
AUTHORS Rolfe,M., Chiu,M.,Isabel., Cottarel,G., Berlin,V., Damagnez,V. and Draetta,G.
TITLE Ubiquitin conjugating enzymes
JOURNAL Patent: US 5968761-A 16 19-OCT-1999;
FEATURES location/Qualifiers
source 1. .476
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 100.0%; Score 32; DB 2; Length 476;
Best Local Similarity 100.0%; Pred. No. 0.0031;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TTATTATAAGTGCTCGCGTGCAGAAACC 32
301 TTATTATAAGTGCTCGCGTGCAGAAACC 332

Db 301 TTATTATAAGTGCTCGCGTGCAGAAACC 332

RESULT 6
AR085249 476 bp DNA linear PAT 01-SEP-2000
LOCUS Sequence 9 from patent US 5981699.
DEFINITION AR085249
ACCESSION AR085249 GI:10012019
VERSION AR085249.1
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 476)
AUTHORS Draetta,G., Rolfe,M. and Eckstein,J.W.
TITLE Human ubiquitin conjugating enzyme
JOURNAL Patent: US 5981699-A 9 09-NOV-1999;
FEATURES location/Qualifiers
source 1. .476
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 100.0%; Score 32; DB 2; Length 476;
Best Local Similarity 100.0%; Pred. No. 0.0031;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TTATTATAAGTGCTCGCGTGCAGAAACC 32
301 TTATTATAAGTGCTCGCGTGCAGAAACC 332

Db 301 TTATTATAAGTGCTCGCGTGCAGAAACC 332

RESULT 7
CS113238 477 bp DNA linear PAT 24-JUN-2005
LOCUS Sequence 15 from Patent WO2005051431.
DEFINITION CS113238
ACCESSION CS113238
VERSION CS113238.1 GI:68224783
KEYWORDS
SOURCE Human papillomavirus - 18
ORGANISM Human papillomavirus - 18
REFERENCE 1
AUTHORS Milner,A.J.
TITLE Colloidal delivery system for biological therapeutic agents
JOURNAL Patent: WO 2005051431-A 15 09-JUN-2005;
FEATURES location/Qualifiers
source 1. .477
/organism="Human papillomavirus - 18"
/mol_type="unassigned DNA"
/db_xref="taxon:337042"

ORIGIN
Query Match 100.0%; Score 32; DB 2; Length 477;
Best Local Similarity 100.0%; Pred. No. 0.0031;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TTATTATAAGTGCTCGCGTGCAGAAACC 32
301 TTATTATAAGTGCTCGCGTGCAGAAACC 332

Db 301 TTATTATAAGTGCTCGCGTGCAGAAACC 332

RESULT 8
PAHPV181 528 bp DNA linear VRL 06-JUL-1989
LOCUS Integrated HPV-18 E6/E7 ORF with DNaseI-hypersensitive site.
DEFINITION X05568
ACCESSION X05568 GI:60880
VERSION X05568.1
KEYWORDS DNase I hypersensitive site; E6 gene; E7 gene; enhancer-like sequence; inverted repeat.
SOURCE Human papillomavirus
ORGANISM Human papillomavirus
REFERENCE 1 (bases 1 to 528)
AUTHORS Lazo,P.A.
TITLE Structure, DNaseI hypersensitivity and expression of integrated papilloma virus in the genome of HeLa cells
JOURNAL Bur. J. Biochem. 165 (2), 393-401 (1987)
COMMENT 2439332
numbering refers to the Xba site in E6 ORF
Data kindly reviewed (12.2.88) by Lazo P. A.
FEATURES location/Qualifiers
source 1. .528
/organism="Human papillomavirus"
/mol_type="genomic DNA"
/db_xref="taxon:10566"
/cell_line="HeLa"
18. .24
/note="enhancer-like sequence 1"
21. .28
/note="inverted repeat A"
complement (34. .42)
/note="enhancer-like sequence 2"
38. .42
/note="inverted repeat B"
46. .53
/note="enhancer-like sequence 3"
77. .81
/note="inverted repeat B"
90. .97
repeat_unit

misc_feature /note="imp.inverted repeat A'"
139..145
/note="enhancer-like sequence 4"
misc_feature 164..171
/note="enhancer-like sequence 5"
misc_feature 250..257
/note="enhancer-like sequence 6"
misc_feature 296..303
/note="enhancer-like sequence 7"
misc_feature 435..442
/note="enhancer-like sequence 8"
misc_feature 445..452
/note="enhancer-like sequence 9"
misc_feature complement(453..460)
/note="enhancer-like sequence 10"
misc_feature complement(460..466)
/note="enhancer-like sequence 11"
misc_feature complement(470..477)
/note="enhancer-like sequence 12"

ORIGIN

Query Match 100.0%; Score 32; DB 10; Length 528;
Best Local Similarity 100.0%; Pred. No. 0.003;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTATTATAAGTGCTGCGGTGCCAGAAACC 32
85 TTATTATAAGTGCTGCGGTGCCAGAAACC 116

RESULT 9
188814 LOCUS 817 bp DNA linear PAT 10-AUG-1998
DEFINITION Sequence 11 from patent US 5719054.
ACCESSION 188814
VERSION 188814.1 GI:3408754
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 817)
AUTHORS Bourneil,M.E., Ingile,S.C. and Munro,A.J.
TITLES Recombinant virus vectors encoding human papillomavirus proteins
JOURNAL Patent: US 5719054-A 11 17-FEB-1998;
FEATURES
source 1..817
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN

Query Match 100.0%; Score 32; DB 2; Length 817;
Best Local Similarity 100.0%; Pred. No. 0.0029;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTATTATAAGTGCTGCGGTGCCAGAAACC 32
306 TTATTATAAGTGCTGCGGTGCCAGAAACC 337

RESULT 10
A98759 LOCUS 837 bp DNA linear PAT 26-JAN-2000
DEFINITION Sequence 20 from Patent WO9910375.
ACCESSION A98759
VERSION A98759.1 GI:6781783
KEYWORDS
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 837)
AUTHORS Lombardo-Benchikh,A. and Bruck,C.
TITLES VACCINE
JOURNAL Patent: WO 9910375-A 20 04-MAR-1999;

FEATURES
source 1..837
/organism="unidentified"
/mol_type="unassigned DNA"
/db_xref="taxon:32644"

ORIGIN

Query Match 100.0%; Score 32; DB 2; Length 837;
Best Local Similarity 100.0%; Pred. No. 0.0029;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTATTATAAGTGCTGCGGTGCCAGAAACC 32
634 TTATTATAAGTGCTGCGGTGCCAGAAACC 665

RESULT 11
BD080320 LOCUS 837 bp DNA linear PAT 27-AUG-2002
DEFINITION Vaccine.
ACCESSION BD080320
VERSION BD080320.1 GI:22625923
KEYWORDS JP 2001513986-A/10.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Homo.
1 (bases 1 to 837)
Bruck,C., Silva,T.C., Delisse,A.M.E.F., Gerard,C.M.G. and
Benchikh,A.L.
Vaccine
Patent: JP 2001513986-A 10 11-SEP-2001;
SMITHKLINE BEECHAM BIOLOGICALS SA
COMMENT OS Homo sapiens (human)
PN JP 2001513986-A/10
PD 11-SEP-2001
PF 17-AUG-1998 JP 2000507701
PR 22-AUG-1997 GB 9717953.5
PI CLAUDINE BRUCK,TERESA CABEZON SILVA,ANNE MARIE EVA FERNANDE
PI DELISSE,
PI CATHERINE MARIE GHISLAINE GERARD,ANGELA LOMBARDO BENCHEIKH PC
C12N15/09,A61K9/107,A61K39/12,A61K39/145,A61K39/39,A61P35/00, PC
A61P37/00,
PC C07K14/025,C07K19/00,C12N1/15,C12N1/19,C12N1/21,C12N5/10, PC
C12P21/02,
PC C12N15/00,C12N5/00
CC Vaccine
FH Key
FT source 1..837
Location/Qualifiers
/organism="Homo sapiens (human)".
1..837
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

FEATURES
source 1..837
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 100.0%; Score 32; DB 2; Length 837;
Best Local Similarity 100.0%; Pred. No. 0.0029;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTATTATAAGTGCTGCGGTGCCAGAAACC 32
634 TTATTATAAGTGCTGCGGTGCCAGAAACC 665

RESULT 12
BD103169 LOCUS 837 bp DNA linear PAT 27-AUG-2002
DEFINITION Vaccine.
ACCESSION BD103169

VERSION BD103169.1 GI:22648743
KEYWORDS JP 2001527091-A/10.
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 837)
AUTHORS Dalemans, W.L.J. and Gerard, C.M.G.
TITLE Vaccine
JOURNAL Patent: JP 2001527091-A 10-25-DEC-2001;
SMITHKLINE BEECHAM BIOLOGICALS SA
COMMENT OS Unidentified
PN JP 2001527091-A/10
PD 25-DEC-2001
PF 18-DEC-1998 JP 2000526542
PR 24-DEC-1997 GB 9727262.9
PI WILFRIED L.J DALEMAN, CATHERINE MARIE GHISLAINE GERARD PC
C07K14/025, A61K39/00, A61K39/12, A61K39/385, A61P35/00, C07K19/00, PC
C12N15/09,
PC Strandedness: Single;
CC Topology: Linear;
CC Vaccine
FH Key
FT source
FEATURES
source 1. .837 Location/Qualifiers
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

ORIGIN
Query Match 100.0%; Score 32; DB 2; Length 837;
Best Local Similarity 100.0%; Pred. No. 0.0029;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TTATTATTAAGGTCCTGCGGTGCCAGAAACC 32
|||||
634 TTATTATTAAGGTCCTGCGGTGCCAGAAACC 665

Db 634 TTATTATTAAGGTCCTGCGGTGCCAGAAACC 665

RESULT 13
ARI83581 837 bp DNA linear PAT 20-APR-2002
LOCUS
DEFINITION Sequence 20 from patent US 6342224.
ACCESSION ARI83581
VERSION ARI83581.1 GI:20227550
KEYWORDS
SOURCE Unknown.
ORGANISM Unidentified.
REFERENCE 1 (bases 1 to 837)
AUTHORS Bruck, C., Silva, T., Gabezon, F., Bernarde Delisse, A.-M., Eva, C., Ghislaine Gerard, C., Marve, A. and Lombardo-Bencheikh, A.
TITLE Recombinant papillomavirus vaccine and method for production and treatment
JOURNAL Patent: US 6342224-A 20-29-JAN-2002;
FEATURES
source 1. .837 Location/Qualifiers
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 100.0%; Score 32; DB 2; Length 837;
Best Local Similarity 100.0%; Pred. No. 0.0029;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TTATTATTAAGGTCCTGCGGTGCCAGAAACC 32
|||||
634 TTATTATTAAGGTCCTGCGGTGCCAGAAACC 665

Db 634 TTATTATTAAGGTCCTGCGGTGCCAGAAACC 665

RESULT 14

AX020943
LOCUS AX020943 837 bp DNA linear PAT 07-SEP-2000
DEFINITION Sequence 20 from Patent WO9933866.
ACCESSION AX020943
VERSION AX020943.1 GI:10044609
KEYWORDS
SOURCE Human papillomavirus
ORGANISM Human papillomavirus
REFERENCE 1 (bases 1 to 837)
AUTHORS Dalemans, W.L. and Gerard, C.M.
TITLE Vaccine
JOURNAL Patent: WO 9933868-A 20-08-JUL-1999;
DALEMAN WILFRIED L.J (BE); SMITHKLINE BEECHAM BIOLOG (BE); GERARD CATHERINE MARIE GHISLAINE (BE)
FEATURES
source 1. .837 Location/Qualifiers
/organism="Human papillomavirus"
/mol_type="unassigned DNA"
/db_xref="taxon:10566"

ORIGIN
Query Match 100.0%; Score 32; DB 2; Length 837;
Best Local Similarity 100.0%; Pred. No. 0.0029;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TTATTATTAAGGTCCTGCGGTGCCAGAAACC 32
|||||
634 TTATTATTAAGGTCCTGCGGTGCCAGAAACC 665

Db 634 TTATTATTAAGGTCCTGCGGTGCCAGAAACC 665

RESULT 15
HPY18491 843 bp DNA linear VRL 18-AUG-1999
LOCUS
DEFINITION Human papillomavirus type 18 E6 and E7 genes.
ACCESSION Y18491
VERSION Y18491.1 GI:5748503
KEYWORDS E6 gene; E6 protein; E7 gene; E7 protein.
SOURCE Human papillomavirus type 18
ORGANISM Human papillomavirus type 18
REFERENCE 1 (bases 1 to 843)
AUTHORS Laasri, M., Gul'ko, L., Vinokurova, S., Kissel'jova, N., Veiko, V. and Kissel'jov, F.
TITLE Cloning of E6 and E7 Genes of Human Papilloma Virus Type 18 and Transformation Potential of E7 Gene and its Mutants
JOURNAL Virus Genes 182, 139-149 (1999)
REFERENCE 2 (bases 1 to 843)
AUTHORS Veiko, V.P.
TITLE Direct Submission
JOURNAL Submitted (02-DEC-1998) V.P. Veiko, Institute of Genetics and Selection of Industrial Microorganisms, Iet Dorozhny proezd, 1, Moscow 113545, RUSSIA
FEATURES
source 1. .843 Location/Qualifiers
/organism="Human papillomavirus type 18"
/mol_type="genomic DNA"
/db_xref="taxon:333761"
40..516
/gene="E6"
40..516
/gene="E6"
/gene="E6"
/codon_start=1
/product="E6 protein"
/protein_id="CAB53096.1"
/db_xref="GI:5748504"
/db_xref="GOA:Q9QNP8"
/db_xref="UniProtKB/TREMBL:Q9QNP8"
/translation="MARFSDPTRRPYLPLPCLTBLNTSLDIDETTCYCTVLELTV
FERAFKDLFVYRDSIIPHAACHKCIDPYSIRIELRHYSVSVGGTLEKLTNTGTYNLL
IRCIKCKPLNPAKLEKLHLEKRRFHKIAGHYRGQCHSCCNRAQERLQRRRETOV"

gene 525..842
/gene="E7"
CDS 525..842
/gene="E7"
/codon_start=1
/product="E7 protein"
/protein_id="CA53097.1"
/db_xref="GI:5748505"
/db_xref="UniProtKB/TrEMBL:O90NP7"
/translation="MHGPKATLQDYLHLHPONEIPVDLLCHROLSDBSEBNDIDGV
NHQHLPARAEPRHTMLCMCKCEARIKLVVSSADDLRAFOQLFLNTLSFVCPWCA
SQO"

ORIGIN

Query Match 100.0%; Score 32; DB 10; Length 843;
Best Local Similarity 100.0%; Pred. No. 0.0028;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTATTATAAGTGCTGCGGTCGCGCAAAACC 32
|||||
340 TTATTATAAGTGCTGCGGTCGCGCAAAACC 371

Db 340 TTATTATAAGTGCTGCGGTCGCGCAAAACC 371

RESULT 16
PPH18A 912 bp DNA linear VRL 02-AUG-1993
LOCUS Human papillomavirus type 18 encoding envelope proteins E6 and E7,
DEFINITION complete cds., and envelope protein E1, 5' end.
ACCESSION M26798.1 GI:333043
VERSION E1 protein; E6 gene; E7 gene; envelope-associated protein.
KEYWORDS Human papillomavirus
SOURCE Human papillomavirus
ORGANISM Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
unclassified Papillomaviridae.
REFERENCE 1 (bases 1 to 912)
Schneider-Gadicko, A. and Schwarz, E.
AUTHORS Different human cervical carcinoma cell lines show similar
TITLE transcritpion patterns of human papillomavirus type 18 early genes
JOURNAL EMBO J. 5 (9), 2285-2292 (1986)
COMMENT Original source text: Human papillomavirus type 18, cDNA to mRNA,
PUBMED from human cervical carcinoma HeLa, C4-1 and SW756 cell lines.
FEATURES
source location/Qualifiers
1..912
/organism="Human papillomavirus"
/mol_type="genomic DNA"
/db_xref="taxon:10566"
12..488
/note="ORF E6"
/codon_start=1
/protein_id="AAA46946.1"
/db_xref="GI:333044"
/translation="MARFSDPTRRPYKLPDLCTELSTLSODIRITCVYCKTVLELTV
FERAFDLPVYVRDSIPHAACHKCIDPYRIELRLRYSDSVGDTLEKLTNGLYML
IRCLRCQKPLNPAEKRLHNEKRFRFNIAGHYGQCHSCCNRAORLORRRTOY"
497..814
/note="ORF E7"
/codon_start=1
/protein_id="AAA46947.1"
/db_xref="GI:333045"
/translation="MHGPKATLQDYLHLHPONEIPVDLLCHROLSDBSEBNDIDGV
NHQHLPARAEPRHTMLCMCKCEARIKLVVSSADDLRAFOQLFLNTLSFVCPWCA
SQO"
821..>912
/note="ORF E1"
/codon_start=1
/protein_id="AAA46948.1"
/db_xref="GI:333046"
/translation="MADPBGTDGEGTGCNGMFVVOALVDKKTGDV"

ORIGIN

Query Match 100.0%; Score 32; DB 10; Length 912;

Best Local Similarity 100.0%; Pred. No. 0.0029;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTATTATAAGTGCTGCGGTCGCGCAAAACC 32
|||||
312 TTATTATAAGTGCTGCGGTCGCGCAAAACC 343

Db 312 TTATTATAAGTGCTGCGGTCGCGCAAAACC 343

RESULT 17
A06328 1000 bp RNA linear PAT 23-JUL-1993
LOCUS Hpv18 mRNA for E6, E7 and E1.
DEFINITION A06328
ACCESSION A06328.1 GI:413673
VERSION E1 gene; E6 gene; E7 gene.
KEYWORDS Human papillomavirus type 18
SOURCE Human papillomavirus type 18
ORGANISM Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
Alphapapillomavirus.
REFERENCE 1 (bases 1 to 1000)
Oltersdorf, T., Roewer, W., Schneider-Gadicko, A., Seedorf, K.,
AUTHORS Duerst, M. and Schwarz, E.
TITLE Expression products of human papilloma virus type 18, antibodies
specific for these proteins, and diagnostic reagents containing
these antibodies or the corresponding DNA
PATENT: EP 0256321-A 5 24-FEB-1988;
BEHRINGWERKE Aktiengesellschaft
location/Qualifiers
1..1000
/organism="Human papillomavirus type 18"
/mol_type="unassigned RNA"
/db_xref="taxon:333761"
100..576
/codon_start=1
/product="E6"
/protein_id="CAA00542.1"
/db_xref="GI:413674"
/db_xref="GOA:P06788"
/db_xref="UniProtKB/Swiss-Prot:P06788"
/translation="MARFSDPTRRPYKLPDLCTELSTLSODIRITCVYCKTVLELTV
FERAFDLPVYVRDSIPHAACHKCIDPYRIELRLRYSDSVGDTLEKLTNGLYML
IRCLRCQKPLNPAEKRLHNEKRFRFNIAGHYGQCHSCCNRAORLORRRTOY"
585..902
/codon_start=1
/product="E7"
/protein_id="CAA00543.1"
/db_xref="GI:413675"
/db_xref="GOA:P06788"
/db_xref="UniProtKB/Swiss-Prot:P06788"
/translation="MHGPKATLQDYLHLHPONEIPVDLLCHROLSDBSEBNDIDGV
NHQHLPARAEPRHTMLCMCKCEARIKLVVSSADDLRAFOQLFLNTLSFVCPWCA
SQO"
909..>1000
/codon_start=1
/product="E1"
/protein_id="CAA00544.1"
/db_xref="GI:413676"
/translation="MADPBGTDGEGTGCNGMFVVOALVDKKTGDV"

ORIGIN

Query Match 100.0%; Score 32; DB 2; Length 1000;
Best Local Similarity 100.0%; Pred. No. 0.0028;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTATTATAAGTGCTGCGGTCGCGCAAAACC 32
|||||
400 TTATTATAAGTGCTGCGGTCGCGCAAAACC 431

Db 400 TTATTATAAGTGCTGCGGTCGCGCAAAACC 431

RESULT 18
A07621 1000 bp RNA linear PAT 22-JUN-1993
LOCUS Amplified spliced mRNA for HPV 18.
DEFINITION

ACCESSION A07621
VERSION A07621.1 GI:413119
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
OTHER SEQUENCES; artificial sequences.
REFERENCE 1 (bases 1 to 1000)
AUTHORS Cerutti, P., Whitcomb, J., Zijlstra, J. and De Villiers, E.M.
TITLE Detection of human papillomavirus DNA in cervix smears
JOURNAL Patent: EP 0373352-A 7 20-JUN-1990;
BEHRINGERWERKE Aktiengesellschaft
FEATURES
source 1. 1000
/organism="synthetic construct"
/mol_type="unassigned RNA"
/db_xref="taxon:32630"
ORIGIN
Query Match 100.0%; Score 32; DB 2; Length 1000;
Best Local Similarity 100.0%; Pred. No. 0.0028;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 TTATTATAAGTGCTCGCGTGCCAGAAACC 32
Db 405 TTATTATAAGTGCTCGCGTGCCAGAAACC 436
RESULT 19
LOCUS A98761 1152 bp DNA linear PAT 26-JAN-2000
DEFINITION Sequence 22 from Patent WO910375.
ACCESSION A98761
VERSION A98761.1 GI:6781784
KEYWORDS
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 1152)
AUTHORS Lombardo-Bencheikh, A. and Bruck, C.
TITLE VACCINE
JOURNAL Patent: WO 9910375-A 22 04-MAR-1999;
LOMBARDO BENCHEIKH ANGELA (BE); SMITHKLINE BEECHAM BIOLOG (BE)
FEATURES
source 1. 1152
/organism="unidentified"
/mol_type="unassigned DNA"
/db_xref="taxon:32644"
ORIGIN
Query Match 100.0%; Score 32; DB 2; Length 1152;
Best Local Similarity 100.0%; Pred. No. 0.0028;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 TTATTATAAGTGCTCGCGTGCCAGAAACC 32
Db 634 TTATTATAAGTGCTCGCGTGCCAGAAACC 665
RESULT 20
LOCUS BD080321 1152 bp DNA linear PAT 27-AUG-2002
DEFINITION Vaccine.
ACCESSION BD080321
VERSION BD080321.1 GI:22625924
KEYWORDS JP 2001513986-A/11.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.
REFERENCE 1 (bases 1 to 1152)
AUTHORS Bruck, C., Silva, T.C., Delisse, A.M.B.F., Gerard, C.M.G. and
Bencheikh, A.L.

TITLE Vaccine
JOURNAL Patent: JP 2001513986-A 11 11-SEP-2001;
SMITHKLINE BEECHAM BIOLOGICALS SA
COMMENT OS Homo sapiens (human)
PN JP 2001513986-A/11
PD 11-SEP-2001
PF 17-AUG-1998 JP 2000507701
PR 22-AUG-1997 GB 9717953.5
PI CLAUDINE BRUCK, TERESA CABEZON SILVA, ANNE MARIE EVA FERNANDE
DELISSE,
PI CATHERINE MARIE GHISLAINE GERARD, ANGELA LOMBARDO BENCHEIKH PC
C12N15/09, A61K39/107, A61K39/12, A61K39/145, A61K39/39, A61P35/00, PC
A61P37/00,
PC C07K14/025, C07K19/00, C12N1/15, C12N1/19, C12N1/21, C12N5/10, PC
C12P21/02,
PC C12N15/00, C12N5/00
CC Vaccine
CC Key
FH Key
FT source 1. 1152
Location/Qualifiers
/organism="Homo sapiens (human)".
FEATURES
source 1. 1152
Location/Qualifiers
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 100.0%; Score 32; DB 2; Length 1152;
Best Local Similarity 100.0%; Pred. No. 0.0028;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 TTATTATAAGTGCTCGCGTGCCAGAAACC 32
Db 634 TTATTATAAGTGCTCGCGTGCCAGAAACC 665
RESULT 21
LOCUS BD103170 1152 bp DNA linear PAT 27-AUG-2002
DEFINITION Vaccine.
ACCESSION BD103170
VERSION BD103170.1 GI:22648744
KEYWORDS JP 2001527091-A/11.
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 1152)
AUTHORS Dalemans, W.L.J. and Gerard, C.M.G.
TITLE VACCINE
JOURNAL Patent: JP 2001527091-A 11 25-DEC-2001;
SMITHKLINE BEECHAM BIOLOGICALS SA
COMMENT OS Unidentified
PN JP 2001527091-A/11
PD 25-DEC-2001
PF 18-DEC-1998 JP 2000526542
PR 24-DEC-1997 GB 9727262.9
PI WILFRIED L J DALEMANS, CATHERINE MARIE GHISLAINE GERARD PC
C07K14/025, A61K39/00, A61K39/12, A61K39/385, A61P35/00, C07K19/00, PC
C12N15/09,
PC C12N15/00
CC Strandedness: Single;
CC Topology: Linear;
CC Vaccine
FH Key
FT source 1. 1152
Location/Qualifiers
/organism="unidentified".
FEATURES
source 1. 1152
Location/Qualifiers
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
ORIGIN

Query Match 100.0%; Score 32; DB 2; Length 1152;
Best Local Similarity 100.0%; Pred. No. 0.0028;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTATTATAAGTGCTGCGGTGCCAGAAACC 32
634 TTATTATAAGTGCTGCGGTGCCAGAAACC 665

RESULT 22
LOCUS AR183582 1152 bp DNA linear PAT 20-APR-2002
DEFINITION Sequence 22 from patent US 6342224.
ACCESSION AR183582
VERSION AR183582.1 GI:20227551
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 1152)
AUTHORS Bruck,C., Silva,T,Cabezon., Fernando Dellesse,A.-M.Eva., Ghislaine Gerard,C.Marie. and Lombardo-Benchikh,A.
TITLE Recombinant papillomavirus vaccine and method for production and treatment
JOURNAL Patent: US 6342224-A 22 29-JAN-2002;
FEATURES
LOCUS 1.1152
Location/Qualifiers
source /organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 100.0%; Score 32; DB 2; Length 1152;
Best Local Similarity 100.0%; Pred. No. 0.0028;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTATTATAAGTGCTGCGGTGCCAGAAACC 32
634 TTATTATAAGTGCTGCGGTGCCAGAAACC 665

RESULT 23
LOCUS AX020945 1152 bp DNA linear PAT 07-SEP-2000
DEFINITION Sequence 22 from Patent WO9333868.
ACCESSION AX020945
VERSION AX020945.1 GI:10044610
KEYWORDS Human papillomavirus
SOURCE Human papillomavirus
ORGANISM Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
REFERENCE 1
AUTHORS Dalemans,W.L. and Gerard,C.M.
TITLE Vaccine
JOURNAL Patent: WO 9333868-A 22 08-JUL-1999;
DALEMAN WILFRIED L J (BE); SMITHKLINE BEECHAM BIOLOG (BE); GERARD CATHERINE MARIE GHISLAI (BE)
FEATURES
LOCUS 1.1152
Location/Qualifiers
source /organism="Human papillomavirus"
/mol_type="unassigned DNA"
/db_xref="taxon:10566"

ORIGIN
Query Match 100.0%; Score 32; DB 2; Length 1152;
Best Local Similarity 100.0%; Pred. No. 0.0028;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTATTATAAGTGCTGCGGTGCCAGAAACC 32
634 TTATTATAAGTGCTGCGGTGCCAGAAACC 665

RESULT 24
LOCUS BD015779 1188 bp DNA linear PAT 27-AUG-2002
DEFINITION Method for detecting human carcinogenic papilloma virus.
ACCESSION BD015779
VERSION BD015779.1 GI:22556916
KEYWORDS JP 2001197894-A/1.
SOURCE Papillomaviridae
ORGANISM Papillomaviridae
REFERENCE 1 (bases 1 to 1188)
AUTHORS Morris,B.J. and Nightingale,B.
TITLE Method for detecting human carcinogenic papilloma virus
JOURNAL Patent: JP 2001197894-A 1 24-JUL-2001;
COMMENT BIOSEARCH INTERNATIONAL PTY LTD
OS Papillomavirus
PN JP 2001197894-A/1
PD 24-JUL-2001 JP 2000379281
PF 13-DEC-2000 JP 2000379281
PR 26-FEB-1987 AU P10559
PI BRYAN JAMES MORRIS, BRYAN NIGHTINGALE
PC C12N15/09,C12Q1/42,C12Q1/68,C12Q1/70//C12Q1/42,C12R1:01, PC
(C12Q1/42,C12R1:93),(C12Q1/68,C12R1:01),(C12Q1/68,C12R1:93), PC
(C12Q1/70,C12R1:93),(C12Q1/70,C12R1:01),C12N15/00 CC Method for
detecting human carcinogenic papilloma virus FH Key
Location/Qualifiers
FT source 1.1188
/organism="Papillomavirus".
FEATURES
LOCUS 1.1188
Location/Qualifiers
source /organism="Papillomaviridae"
/mol_type="genomic DNA"
/db_xref="taxon:151340"

ORIGIN
Query Match 100.0%; Score 32; DB 2; Length 1188;
Best Local Similarity 100.0%; Pred. No. 0.0028;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTATTATAAGTGCTGCGGTGCCAGAAACC 32
610 TTATTATAAGTGCTGCGGTGCCAGAAACC 641

RESULT 25
LOCUS B54157 1188 bp DNA linear PAT 31-JAN-2002
DEFINITION Method for detecting human carcinogenic papilloma virus HPV18.
ACCESSION B54157
VERSION B54157.1 GI:18628305
KEYWORDS JP 2000189200-A/1.
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 1188)
AUTHORS James,M.B. and Bryan,N.
TITLE Method for detecting human carcinogenic papilloma virus HPV18
JOURNAL Patent: JP 2000189200-A 1 11-JUL-2000;
COMMENT BIOSEARCH INTERNATIONAL PTY LTD
OS Papilloma virus
PN JP 2000189200-A/1
PD 11-JUL-2000 JP 2000043673
PF 21-FEB-2000 JP 2000043673
PR 26-FEB-1987 AU P10559
PI MAURICE BRYAN JAMES, NIGHTINGALE BRYAN
PC C12Q1/68,C12N15/09,C12Q1/70//C12Q1/70,C12R1:91) CC
FH Key Location/Qualifiers
FT source 1.1188
/organism="Papilloma virus".
FEATURES
LOCUS 1.1188
Location/Qualifiers
source /organism="unidentified"
/mol_type="genomic DNA"

ORIGIN /db_xref="taxon:32644"

Query Match 100.0%; Score 32; DB 2; Length 1188;
Best Local Similarity 100.0%; Pred. No. 0.0028;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTATTATAAGTGCCTGCGTCCAGAAACC 32
|||||
610 TTATTATAAGTGCCTGCGTCCAGAAACC 641

RESULT 26
PARHPV6 1188 bp DNA linear VRL 12-SEP-1993
LOCUS Human papilloma virus (HPV-18) E6 ORF.
X04354
ACCESSION X04354.1 GI:60995
VERSION transforming capacity; unidentified reading frame.
KEYWORDS Human papillomavirus
SOURCE Human papillomavirus
ORGANISM Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
unclassified Papillomaviridae.
1 (bases 1 to 1188)
Matishevskii, G., Banks, L., Wu-Liao, J., Spence, P., Pim, D. and Crawford, L.
The expression of human papillomavirus type 18 E6 protein in bacteria and the production of anti-E6 antibodies
J. Gen. Virol. 67 (Pt 9), 1909-1916 (1986)
3018129

FEATURES
source location/Qualifiers
1..1188
/organism="Human papillomavirus"
/mol_type="genomic DNA"
/strain="type 18"
/db_xref="taxon:10566"
/tissue type="cervical carcinoma"
310..786
/note="unnamed protein product; E6 ORF (aa 1-158)"
/codon_start=1
/protein_id="CA027879.1"
/db_xref="GI:60996"
/db_xref="GOA:P06463"
/translation="MARFEDPTRRPYKLPDLCTELNLSDDIETCYCKTVLETEV
FEFAFKDLFVYRDSIPHAACHCICIDYTSIRLRIHYSVSDYDTLEKLTNTGLYNL
IRCLRCOKPLNPAEKLRHLEKRRFHNIAHYRGQCHSCCNRAQRERLQRRRTQV"

CDS
100.0%; Score 32; DB 10; Length 1188;
Best Local Similarity 100.0%; Pred. No. 0.0028;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

ORIGIN

Query Match 100.0%; Score 32; DB 10; Length 1188;
Best Local Similarity 100.0%; Pred. No. 0.0028;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTATTATAAGTGCCTGCGTCCAGAAACC 32
|||||
610 TTATTATAAGTGCCTGCGTCCAGAAACC 641

RESULT 27
A06324 1750 bp DNA linear PAT 23-JUL-1993
LOCUS Hpv18 genes for E6, E7, and E1.
X06324.1 GI:413669
VERSION E1 gene; E6 gene; E7 gene.
KEYWORDS Human papillomavirus type 18
SOURCE Human papillomavirus type 18
ORGANISM Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
unclassified Papillomaviridae.
1 (bases 1 to 1750)
Oltersdorf, T., Roewekamp, W., Schneider-Gaeddicke, A., Seedorf, K.,
Duerst, M. and Schwarz, E.
Expression products of human papilloma virus type 18, antibodies

specific for these proteins, and diagnostic reagents containing these antibodies or the corresponding DNA
Patent: EP 0256321-A 1 24-FEB-1988;
BEHRINGWERKE Aktiengesellschaft
Location/Qualifiers
1..1750
/organism="Human papillomavirus type 18"
/mol_type="unassigned DNA"
/db_xref="taxon:333761"
118..594
/codon_start=1
/product="E6"
/protein_id="CA00539.1"
/db_xref="GI:413670"
/db_xref="GOA:P06463"
/db_xref="UniProtKB/Swiss-Prot:P06463"
/translation="MARFEDPTRRPYKLPDLCTELNLSDDIETCYCKTVLETEV
FEFAFKDLFVYRDSIPHAACHCICIDYTSIRLRIHYSVSDYDTLEKLTNTGLYNL
IRCLRCOKPLNPAEKLRHLEKRRFHNIAHYRGQCHSCCNRAQRERLQRRRTQV"
603..920
/codon_start=1
/product="E7"
/protein_id="CA00540.1"
/db_xref="GI:413671"
/db_xref="GOA:P06788"
/db_xref="UniProtKB/Swiss-Prot:P06788"
/translation="MGPKATLDIVLHSEVQNBIFPDLICHELQSLSEBENDEIDGV
NHQHLPRARRPQRHTLWCCKCEARLIELVESSADLRAFOQLFNTLSFVCPWCA
SQO"
927..1750
/codon_start=1
/product="E1"
/protein_id="CA00541.1"
/db_xref="GI:413672"
/db_xref="GOA:Q84181"
/db_xref="UniProtKB/TrEMBL:Q84181"
/translation="MADPEGTDEGTGNGMFPYQAIIVDKKTGDIIVSDDEENATDNG
SDMVDPIIDQGTFCBOALETAALFHAQSVHDAQVLAHLKRPAGGSTENSPISGR
LEVDLTSPRLQGISLNSGCKAKRRLFTISDGYGSEVRAIQIQTTHGREGANVC
SGGSTALIDNGTEGNNSSVDGTSNINENVPQCTIAQKDLKRNNOGAMLAIV
KDTYGVSTFDLVNPFKSDKTYCTDWVTAIFGVNPTIABGPKTLIQEPLIYAHIOCL"

ORIGIN

Query Match 100.0%; Score 32; DB 2; Length 1750;
Best Local Similarity 100.0%; Pred. No. 0.0027;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTATTATAAGTGCCTGCGTCCAGAAACC 32
|||||
418 TTATTATAAGTGCCTGCGTCCAGAAACC 449

RESULT 28
PARHPV18E 1750 bp DNA linear VRL 12-SEP-1993
LOCUS Human papilloma virus type 18 (HPV 18) DNA for early region with
proteins E6, E7, and E1.
X04773
ACCESSION X04773.1 GI:60876
VERSION unidentified reading frame.
KEYWORDS Human papillomavirus
SOURCE Human papillomavirus
ORGANISM Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
unclassified Papillomaviridae.
1 (bases 1 to 1750)
Seedorf, K., Oltersdorf, T., Krammer, G. and Roewekamp, W.
Identification of early proteins of the human papilloma viruses
type 16 (HPV 16) and type 18 (HPV 18) in cervical carcinoma cells
EMBO J. 6 (1), 139-144 (1987)
3034571
Data kindly reviewed (31-OCT-1987) by SEEDORF K.
Location/Qualifiers
1..1750

```
/organism="Human papillomavirus"
/mol_type="genomic DNA"
/db_xref="taxon:10566"
39..43
/notes="pot. TATA box"
86..90
/notes="pot. TATA box"
118..594
/notes="unnamed protein product; E6 protein (AA 1-158)"
/codon_start=1
/protein_id="CAA28466.1"
/db_xref="GI:60877"
/db_xref="GOA:P06463"
/db_xref="UniProtKB/Swiss-Prot:P06463"
/translation="MARFDPTRPRPKLPDLCTELNTSLQDIEITCYVCKTVELTEV
FEAFKDLFVYVRDSIIPHACHKCIDFYRSIRLRHYSDSVGDTEKLTNTGLYNLL
IRCLRQCKPLNPAEKRLHNEKRRFNINAGHYGQCHSCNRRARORLQRRRTQV"
603..920
/notes="unnamed protein product; E7 protein (AA 1-105)"
/codon_start=1
/protein_id="CAA28467.1"
/db_xref="GI:60878"
/db_xref="GOA:P06788"
/db_xref="UniProtKB/Swiss-Prot:P06788"
/translation="MHGPKATLQDIVLHLEPONEIPVDLLCHEQLSDSEENDEIDGV
NHQHLPRARAEPRHMTLQWCKCKEARIKLVVSSADILRAFOQLFLNTLSFVCPWCA
SQ0"
927..>1750
/notes="unnamed protein product; E1 protein (AA 1-275)"
(1750 is 2nd base in codon)
/codon_start=1
/protein_id="CAA28468.1"
/db_xref="GI:60879"
/db_xref="GOA:Q84181"
/db_xref="UniProtKB/Trembl:Q84181"
/translation="MADPEGTDEGTGCGNMFYQAIIVDKTGDVLSDEDEENATDTG
SDWDFIDTQGFCEQAELETAQALPHAQEVHNDADVLHVKRFPAGSGSTENSLPISR
LEVDTELSRLQELISNGQKAKRRLFTISDYGCSRYEATQIQVTTNGEAGNVC
SGSSTALINDGTGSGNNSVDGTSNSENIVPOCTIAQKDLKYNNGQGMALVF
KDTYGSFTDLVNFKSDKTYCTDWYTAIFGVNPTIABGPKTLIQPFIYAH1QCL"

ORIGIN
Query Match 100.0%; Score 32; DB 10; Length 1750;
Best Local Similarity 100.0%; Pred. No. 0.0027;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTATTAATAAGTGCGTGGCCAGAAACC 32
|||||
418 TTATTAATAAGTGCGTGGCCAGAAACC 449

RESULT 29
HIMHELB 3135 bp mRNA linear VRL 06-DEC-1999
LOCUS Human papillomavirus type 18 proteins E6 and E7 mRNA, complete
DEFINITION
M20325.1 GI:183936
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS Inagaki,Y., Taunokawa,Y., Takebe,N., Nawa,H., Nakanishi,S.,
Terada,H. and Sugimura,T.
Tetrad, nucleotide sequences of cDNAs for human papillomavirus type 18
transcripts in HeLa cells
JOURNAL J. Virol. 62 (5), 1640-1646 (1988)
PUBMED 2833614
COMMENT The integrated viral sequence starts at position 1.
FEATURES
source 1..2393
```

```
/organism="Human papillomavirus type 18"
/mol_type="mRNA"
/db_xref="taxon:333761"
/map="16 bp upstream of BamHI site"
/clone="123"
/cell_line="HeLa; SKG-[I,II]"
/focus
2394..3135
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
1..477
/codon_start=1
/product="protein E6"
/protein_id="AA99514.1"
/db_xref="GI:306836"
/translation="MARFDPTRPRPKLPDLCTELNTSLQDIEITCYVCKTVELTEV
FEAFKDLFVYVRDSIIPHACHKCIDFYRSIRLRHYSDSVGDTEKLTNTGLYNLL
IRCLRQCKPLNPAEKRLHNEKRRFNINAGHYGQCHSCNRRARORLQRRRTQV"
486..803
/codon_start=1
/product="protein E7"
/protein_id="AA99515.1"
/db_xref="GI:306837"
/translation="MHGPKATLQDIVLHLEPONEIPVDLLCHEQLSDSEENDEIDGV
NHQHLPRARAEPRHMTLQWCKCKEARIKLVVSSADILRAFOQLFLNTLSFVCPWCA
SQ0"
810..2396
/codon_start=1
/product="protein E1"
/protein_id="AA99516.1"
/db_xref="GI:306838"
/translation="MADPEGTDEGTGCGNMFYQAIIVDKTGDVLSDEDEENATDTG
SDWDFIDTQGFCEQAELETAQALPHAQEVHNDADVLHVKRFPAGSGSTENSLPISR
LEVDTELSRLQELISNGQKAKRRLFTISDYGCSRYEATQIQVTTNGEAGNVC
SGSSTALINDGTGSGNNSVDGTSNSENIVPOCTIAQKDLKYNNGQGMALVF
KDTYGSFTDLVNFKSDKTYCTDWYTAIFGVNPTIABGPKTLIQPFIYAH1QCLDC
KNGVLDGPDPENFORLTLTIOHGIDSDNFDLSBQVQWAFNDELDSQMAFEYALADS
NSAAAFLESNQAKYLLKDCATWCKHYRAOKROMSQMIRPGSKIDBGQMRPIV
QPLRYQIRPITTLGALXSFLEKTPKKNCLVFCGPANTGSKSYGMSFIHF1QCAV1SF
VNSTSHFMLEPDLTIRVAML"

ORIGIN
Query Match 100.0%; Score 32; DB 10; Length 3135;
Best Local Similarity 100.0%; Pred. No. 0.0025;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTATTAATAAGTGCGTGGCCAGAAACC 32
|||||
301 TTATTAATAAGTGCGTGGCCAGAAACC 332

RESULT 30
HP089349 5210 bp DNA linear VRL 02-DEC-1999
LOCUS Human papillomavirus type 18 variant, partial sequence.
DEFINITION
U89349.1 GI:2052453
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS Schneider-Gedicks,A. and Schwarz,E.
TITLE Different human cervical carcinoma cell lines show similar
transcription patterns of human papillomavirus type 18 early genes
JOURNAL EMBO J. 5 (9), 2285-2292 (1986)
PUBMED 3023067
COMMENT 2 (bases 1 to 5210)
AUTHORS Cole,S.T. and Dancos,O.
TITLE Nucleotide sequence and comparative analysis of the human
```


RESULT 32
AY62282
LOCUS AY62282 7857 bp DNA circular VRL 06-DEC-2005
DEFINITION Human papillomavirus type 18 complete sequence.
ACCESSION AY62282
VERSION AY62282.1 GI:30172004
KEYWORDS
SOURCE
ORGANISM Human papillomavirus type 18
Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
Alphapapillomavirus.
REFERENCE
AUTHORS 1 (bases 1 to 7857)
Narechania, A., Chen, Z., Desalle, R. and Burk, R.D.
TITLE Phylogenetic incongruence among Oncogenic Genital Alpha Human
Papillomaviruses
J. VIROL. 79 (24), 15503-15510 (2005)
JOURNAL
PUBMED 16306621
REFERENCE 2 (bases 1 to 7857)
Chen, Z. and Burk, R.D.
AUTHORS The newly modified full genome sequence of HPV18 prototype (Cole,
TITLE 1987), with E6, E7, E1, E2, E4, E5, L2 and L1 ORFs
Unpublished
3 (bases 1 to 7857)
Burk, R.D.
JOURNAL Direct Submission
AUTHORS Submitted (25-MAR-2003) Microbiology and Immunology, Albert
TITLE Einstein College of Medicine of Yeshiva University, 1300 Morris
Park Ave., Ullmann 515, Bronx, NY 10461, USA
FEATURES
source
1. .7857
/organism="Human papillomavirus type 18"
/mol_type="genomic DNA"
/db_xref="taxon:333761"
105. .581
/gene="E6"
105. .581
/gene="E6"
/codon_start=1
/product="E6 protein"
/protein_id="AAP20594.1"
/db_xref="GI:30172005"
/translation="MARFEDPTRRPKYLPDLCTELTSLQDIEITCYCKTLELTV
FEFAFDLFFVYRDSIPHACHKIDIFYSRIRLRYSDVSDGTLEKLTNTGLVNL
IRCLRCOKPLNPAKELRLNEKRRFHNIAHYRGQSCCNRRARQERLQRRRTQY"
590. .907
/gene="E7"
590. .907
/gene="E7"
/codon_start=1
/product="E7 protein"
/protein_id="AAP20595.1"
/db_xref="GI:30172006"
/translation="MHGPKATLDIYLHLEPQNEIRPYVDLICHQQLSDSENEDELQV
NHQHLPARRAEPQRHTLMCMCKCEARIELIVSSADLRAFOQLFLNTLSFVPCWCA
SQQ"
914. .2887
/gene="E1"
914. .2887
/gene="E1"
/codon_start=1
/product="E1 protein"
/protein_id="AAP20596.1"
/db_xref="GI:30172007"
/translation="MADPSTDSRGTCGNCWFYVQALVTKKTDVTSDDDENATDNG
SDNVDFLDIGTFCROAELSTQALFPAQSVHNDQVHLVLRKRFAGSTENSPLGSR
LEVDTELSPRLQELSNISGQKAKRRLLTISDYGCEVEATQIQVTTNGEAGLVAV
SGSTELADNGTEGNNSSVDGSDNSININVAPOCTIAQKDLKYNNGQAGLVAVF
KQVYGSFTDLVNFKSDKTTCTDMVTAIRGVNPTIAEGFKTLIOPTILVAVHICLDC
KOTVGLIALLRYCGSKRLTVAGLSTLLHVPKCMLOPKLRSSVAAALVYRTGIS
NISEWADPTENQRLITIOHGIDNSFDLSERVQARFDELEIDESMAEYVALADS
NSAAAFKSNCAKTLKDCATCKHYRRAQKQOMNSQWIRFRCSTIDEGQMRPIV
QPLRYQOIEFTFLGALKSPFKGTPKKNCILVFCGPANTGKSYFGMSPIHFIOGAIVSF

gene
CDS
2817. .3914
/gene="E2"
2817. .3914
/gene="E2"
/codon_start=1
/product="E2 protein"
/protein_id="AAP20597.1"
/db_xref="GI:30172008"
/translation="MQTPKATLSERISALQDKIIDRYENDSKDIDSCIQTYOQILREN
ALFFARHEGIQTLNQQVPAVNISSKAKHAIELQMALQGLAQSAVKEDWLTQDTC
EELMNTPEPHCKGGQGVQVYFDGKDKCMKTVMADSVYMTDAGTMDKTAICVSR
GLYVYEGNTFYIEFKSCREKYGNTGWEHGNVINDCNSMCTSDDTVSATQLV
KOLQHPSPYSSTVSVGTAKTGQTSAAIRPGHCGAARQCGAPVPLGAAPPTGN
KRRKCSGNTPTPIIHKGDRNLKCLRYLRHSDHYRDISSTHMTAGNERTGLIT
VYHSETQRKEFLNTVAIPDSVQILVGYTM"
3418. .3684
/gene="E4"
3418. .3684
/gene="E4"
/codon_start=1
/product="E4 protein"
/protein_id="AAP20598.1"
/db_xref="GI:30172009"
/translation="MTLCAPVTRYRPLSLNLSYSTPHRIAPCPMAPOREPTARR
LHDDLTVDSSRSSIYDLSHRSVQLHQTATKQNSVVTTLRL"
3936. .4157
/gene="E5"
3936. .4157
/gene="E5"
/codon_start=1
/product="E5 protein"
/protein_id="AAP20599.1"
/db_xref="GI:30172010"
/translation="MLSLIFLFCFCVCMYCCHVPLLPVSCMCAYAVLVFYIYVIT
SPATAPTVYVFCFLPMLLHIAISIQ"
4244. .5632
/gene="L2"
4244. .5632
/gene="L2"
/codon_start=1
/product="L2 protein"
/protein_id="AAP20600.1"
/db_xref="GI:30172011"
/translation="MYSHRAARRKRASVTDLYTCKOSGTCPPDVVPKYEGTTLADI
LQMSLGITFLGIGTGTSGTGRCQYILPGRSNTVVDVGPTRPVVIBVGPDTDS
ITVLLRDSVVTSGARPTFTGTSGDITSAGTTTAVLDTITPSISVSISTTNPTNP
AFSDP8IIEVPQGEVAGNVFVGTPTSGTHGYEILPQTFASGTESEPISTPLRTV
RRVAGPRLYRAVQGVANPEFLTPS8LITVDNPAFRVHDTTLTFDRSDVDP8DF
MDIIRLRPALTRSGTVRFSRLGQATVFTSGTQIGARVHHDISPLAP8EYIE
LDIPLVAT8DNLDLFDIYADMDPAPV8RSTSG8FAPFK8PTISSASS8NTVPLT
SSMDVYVITGP8DITLTP8TSS8V8I8V8PT8AS8QYIGI8HTHY8L8V8I8Y8I8K8KR
V8I8F8AD8G8V8A8"
5430. .7136
/gene="L1"
5430. .7136
/gene="L1"
/codon_start=1
/product="L1 protein"
/protein_id="AAP20601.1"
/db_xref="GI:30172012"
/translation="MCLTYRVLILYHLLPLVGLYHPRLPLHSILLYVWHIILCGH
YIIFLRNNTVPIFLQMLMRPSDNTVLLP8SVARVNTDVTYRTSYFYAG8SR
LITVGPYRVPVAGG8KODIPKYSAYQVRVPVQVLPDRK8GLPT8SYN8PT8QV
MLCAGVEIG8QPLVGL8GHPYTNKLDIT8SH8A8L8SV8E8V8DN8V8DY8Q8T8Q8C
IICAGVAIL8B8H8AK8G8ACK8SRPL8Q8D8C8P8LE8L8K8N8V8LED8G8DV8DY8G8M8D8ST8LD8
TKCEV8LD8COSI8CKT8PY8LQ8MS8AD8Y8G8MF8CL8RB8OL8F8AH8F8NN8AG8T8M8D8V8Q8
SLYIK8V8K8AR8AS8CV8Y8PS8PG8S8I8V8SD8Q8L8FN8PY8E8L8H8K8Q8H8NN8Q8M8Q8UL8F8V
SY8D8T8R8NT8L8T8CA8T8Q8BV8G8Q8V8AT8K8Q8Y88V8H8E8I8D8Q8L8CTI8T8L8D8V8M8
SY8HS8NN8S8LE8DM8NG8V8PP8T8S8L8V8D8T8R8F8Q8S8A8I8T8Q8D8A8A8E8ND8P8D8K8E8
M8ND8L8K8E8S8LD8ID8Y8PL8G8K8FL8VQ8GL8R8K8R8P8T8IG8R8K8S8AP8A8T8TS8K8AK8K8V8RA

ORIGIN RK

Query Match 100.0%; Score 32; DB 10; Length 7857;
Best Local Similarity 100.0%; Pred. No. 0.0023;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 TTATTATAAGGTGCTGGCGGTGCAGAAACC 32
|||||
405 TTATTATAAGGTGCTGGCGGTGCAGAAACC 436

RESULT 33
PAPPV18 7857 bp DNA linear VRL 18-APR-2005
LOCUS Human papillomavirus type 18 E6, E7, E1, E2, E4, E5, E1 & E2 genes.
ACCESSION X05015.1 GI:60975
VERSION X05015.1 GI:60975
KEYWORDS E1 gene; E2 gene; E5 gene; E6 gene; E7 gene; L1 gene; L2 gene.
SOURCE Human papillomavirus type 18
ORGANISM Human papillomavirus type 18
Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
Alphapapillomavirus.
1
REFERENCE
AUTHORS Cole, S.T. and Danos, O.
TITLE Nucleotide sequence and comparative analysis of the human papillomavirus type 18 genome. Phylogeny of papillomaviruses and repeated structure of the E6 and E7 gene products
JOURNAL J. Mol. Biol. 193 (4), 599-608 (1987)
COMMENT 3039146
FEATURES
source Data kindly reviewed (14-AUG-1987) by Danos O.
location/Qualifiers
1..7857
/organism="Human papillomavirus type 18"
/mol_type="genomic DNA"
/db_xref="taxon:333761"
105..581
/gene="E6"
105..581
/gene="E6"
/codon_start=1
/product="E6 protein"
/protein_id="CA28664.1"
/db_xref="GI:60976"
/db_xref="GOA:P06463"
/db_xref="InterPro:IPR001334"
/db_xref="UniProtKB/Swiss-Prot:P06463"
/translation="MARFEDPTRRKYLPDICTELANTSLQDIETCYCKTVLETEV
FEFAFKDLFVYRDSISIPHACIKCIDFYSLRIELRHYSDSYGGTLEKLTNTGLYNTL
IRCLRCKPLPAKLAHLNKKRFNHIAGHYRQCHSCCNRARQERLQRRRTQV"
590..907
/gene="E7"
590..907
/gene="E7"
/codon_start=1
/product="E7 protein"
/protein_id="CA28665.1"
/db_xref="GI:60977"
/db_xref="GOA:P06788"
/db_xref="UniProtKB/Swiss-Prot:P06788"
/translation="MHGPKATLQDVLIVLHPQNEIPVDLCHQQLSDSEENDEIDGV
NHOHLPARBAEPQHTMLCMCKCEARIELVVESSADDLRAFOQLPLNTLSFVPCWCA
SQQ"
914..2887
/gene="E1"
914..2887
/gene="E1"
/codon_start=1
/product="E1 protein"
/protein_id="CA28666.1"
/db_xref="GI:60978"
/db_xref="GOA:P06789"
/db_xref="InterPro:IPR001177"

gene
CDS
/db_xref="UniProtKB/Swiss-Prot:P06789"
/translation="MADPEGDGEGTGCNGMFYVQALVDDKKTGDI SDDDEBNATDNG
SDMDVFDIQGTGPEQAELEFPAQLPFAQGVHQAQVLAHLVKRRFAGSGTENSILGER
LKYDTERSPRLQELSLNSGOKAKRRLPTISDSYGGSEVQAQTOYTNGEHGNG
SGGSTAIDNGTEGNSVYDGSNENIENVPQCTTAQKDLKLNKKNGANLAVP
KDYGLSFDTLVNRFKSDKTTCTDWTALIGVNPVLAEGFKTLIQPILYAHILCLDC
KMGVLIALLRLKYCKGSRLLVAKGLSTLHVEPCMLIOPKLRSSVAALVYRTGIS
NISEVMDTPEMIQRLTIIOHGLDSDNFDSMVFQAFNDELTDSEMAFEYALDAS
NSNAAFLKSNCAKYLKDCATMCKHYRBAOKRQMNNSOMIRPCSKRIDEGDMRPV
QPLRYQDIERITFGLAKSLFKGTPKRCCLVFCGPATGKSYFEMSPHIFQGVISF
VNSTSHFMLEPLDTKYAMLDDATTTCTTTFDITMRNALDGNPISIDKHKPLQLQKC
PPIILTTNIHPAKDNRPVYLRSLTTFEPNAPFPDKNQNVYINDKMKCFBERTW
SRDLHEEEDADTEGNPFGTFKLRAQGNHRPL"
2817..3914
/gene="E2"
2817..3914
/gene="E2"
/codon_start=1
/product="E2 protein"
/protein_id="CA28667.1"
/db_xref="GI:60979"
/db_xref="GOA:P06790"
/db_xref="UniProtKB/Swiss-Prot:P06790"
/translation="MQPTKETSRLSCVQDKIIDHYENDSKDIDSOIQVQLRMEN
AIFPAAREHGIQTLHQVPAYNISKSAKHAILEMALQGLAQSRYKTEBDMLQDTC
ELIEMTEPHCCKKGGQTOVOVYPDNKNCMTYAMDSVYMTDAGTMDXTARSHR
GLVYKSGYNTFYIFRSBCKRYKNTGWRHFNANNVINDCSNCSSTDDVSAQLV
KQLOHTSPYSSTVSQVATYGTQTSATRGHGLBKQKCGVNPVLAATPGNN
KRLDCSGNTPTPIHLGDRNSLKLCLRYLRKSKHDYRDISSTWMTGAGNEKGIILT
VYHSEYQRTKFLTVAIPDSVQILVGMVTM"
3418..3684
/gene="E4"
3418..3684
/gene="E4"
/codon_start=1
/product="E4 protein"
/protein_id="CA28668.1"
/db_xref="GI:60980"
/db_xref="InterPro:IPR003861"
/db_xref="UniProtKB/Swiss-Prot:P06791"
/translation="MTLCAPVTTRYPLSLNSISTPHRIIPAPCPMAFORPTARRR
LHDLDTVDSRRSIVDLSTHFSVOLHQLATYDGNVVTLRLL"
3936..4157
/gene="E5"
3936..4157
/gene="E5"
/codon_start=1
/product="E5 protein"
/protein_id="CA28669.1"
/db_xref="GI:60981"
/db_xref="InterPro:IPR004270"
/db_xref="UniProtKB/Swiss-Prot:P06792"
/translation="MLSLIFLFCFCVMYCVHPLLPSCMCAVAVLVFVYIVIT
SPATAFVYVFCFLPMLLHLHIALSLQ"
4244..5632
/gene="L2"
4244..5632
/gene="L2"
/codon_start=1
/product="L2 protein"
/protein_id="CA28670.1"
/db_xref="GI:60982"
/db_xref="GOA:P06793"
/db_xref="InterPro:IPR000784"
/db_xref="UniProtKB/Swiss-Prot:P06793"
/translation="WVSHRAARARAVTLKYCKCKSGTCCPPDVVKBEGTTLADKI
LQMSLIGIFLGGLGIGTSGGTGRTGYIPLGSGSNVVDVQTPPVIVIBVGTDS
IVTLIEDSVVTSGARPTFTGSGPTPLGATGTPVLDVTPSTVSIETNFNP
AFSDPSTIIVQNGEVANGVFNCPNPGSGTHGYEILPIOTAFSGTGEPSSTPLPV
RVAAGPRLYRAIQGVSVANPEPLTRSSLIITYNAPFVDTTLITDPRRDVDSDE
MDIILRLRPALTSRGTVPFRSLGQRAIWFRTSGTQIGARHFTHDSPLASPEYIE
LQPLVSAITENDNLEIDYADMDPAVPVPSSTTSFAPFKYSPITSSASSVSNVPLT
SSMDVYVYTGPDITLIPSTSVWPIVSPTABASTOYIGIHGTHVYLMPLYPFKRKA

Gene
5430..7136
/gene="L1"
5430..7136
/gene="L1"
/codon_start=1
/product="L1 protein"
/protein_id="CAA28671.1"
/db_xref="GI:60983"
/db_xref="GOA:P06794"
/db_xref="InterPro:IPR002210"
/db_xref="InterPro:IPR008975"
/db_xref="UniProtKB/Swiss-Prot:P06794"
/translation="MCLYKRVLLIHYLLPLVGLYHPPLPLSHLVVWHIIICSH
YILFLRNVAVFPILQMALMRPSDNTVLPSPVARVNVTDVYTPTSIFHYAGSSR
LITVGNPVPVPAGGKNKODIPKVASVYRFRVQLPDPKFKLPDTSIYNPTORLV
MACGVIRGPGPLGVGLSGHPFYNNKLDTPRESSHAATSNVSEVRNVSVDYQOTLC
ILGCAPIGIRHMAKTRACKRPLSQGDCPPLKNTKLVLEDGWDVYGYAMDSTIYD
TKCEVPLDICOQSICTKPYDLYQMSADPGDSMFCLRRBQLPAHFNRAQTMQDTPQ
SLYIKGTGMPASPGSCVSPSPSGIVTSPSQLFNRPYMLHKAQGHNGVCMNQLFV
TVVDTPSTLTICASTQSPVPGQYDAIKFYRSHVEYDLOFIFOLCTITLALVM
SYIHSNNSSILIEDMNFNGVPPPTSLVDYRFVQSVAILTCQMAAPAEKNDPVDKLF
KNVDLKEKPSLDIDQYPLGRKFLVQAGLRKRPITGRKKSAPSATTSKRAKVRVRA
RK"

ORIGIN

Query Match 100.0%; Score 32; DB 10; Length 7857;
Best Local Similarity 100.0%; Pred. No. 0.0023;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTATTATAAGTGCTGCGGTCCAGAAACC 32
|||
405 TTATTATAAGTGCTGCGGTCCAGAAACC 436

RESULT 34
HPV45E67 910 bp DNA linear VRL 20-MAY-1997
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

HPV45E67 910 bp DNA linear VRL 20-MAY-1997
Human papillomavirus type 45, E6 and E7 genes.
Y13218.1 GI:2113839
E6 gene; E7 gene.
Human papillomavirus type 45
Human papillomavirus type 45
viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
Alphapapillomavirus.

REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

Saxte-Garau,X., Favre,M., Couturier,J. and Orth,G.
Distinct patterns of alteration of myc genes associated with
integration of HPV16 or HPV45 DNA in two genital tumors
J. Gen. Virol.
2 (bases 1 to 910)
Favre,M.
Direct Submission
Submitted (09-MAY-1997) M. Favre, Institut Pasteur, Virology, 25
Rue du Dr Roux, 75015-Paris, FRANCE
location/Qualifiers
1..910
/organism="Human papillomavirus type 45"
/mol_type="genomic DNA"
/isolate="TC4 HPV45 variant"
/db_xref="taxon:10593"
102..578
/gene="E6"
102..578
/gene="E6"
102..578
/gene="E6"
/codon_start=1
/product="oncoprotein E6"
/protein_id="CAA73660.1"
/db_xref="GI:2113840"
/db_xref="GOA:O10608"
/db_xref="UniProtKB/TrEMBL:O10608"
/translation="MARFDDPTQRPYKLPDLCTELNTSLQVSIACVYCKATLERTEV

FEATURES
source
gene
CDS

Gene
587..907
/gene="E7"
587..907
/gene="E7"
/codon_start=1
/product="oncoprotein E7"
/protein_id="CAA73661.1"
/db_xref="GI:2113841"
/db_xref="UniProtKB/TrEMBL:O10609"
/translation="MGRPRATLDEIYLAHPQNELPVDLLCYEQLSSESEENDEADG
VSHAQIPARRAEPQRKILCVCKCKGRILETVSSADLRTLQQLFLSTLSVCPWC
ATNQ"

ORIGIN

Query Match 95.0%; Score 30.4; DB 10; Length 910;
Best Local Similarity 96.9%; Pred. No. 0.015;
Matches 31; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TTATTATAAGTGCTGCGGTCCAGAAACC 32
|||
402 TTATTATAAGTGCTGCGGTCCAGAAACC 433

RESULT 35
PPH45E67A 1843 bp DNA linear VRL 02-AUG-1993
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

PPH45E67A 1843 bp DNA linear VRL 02-AUG-1993
Human papillomavirus type 45 E6 and E7 genes, complete cds.
M38198.1 GI:333059
Human papillomavirus
Human papillomavirus
viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
unclassified Papillomaviridae.
1 (bases 1 to 1843)
Kaplan,U.B. and Burk,R.D.
Evolution of the Papillomaviruses
Unpublished (1990)
Original source text: Human papillomavirus (type 45) DNA.
Draft entry and computer-readable sequence for [1] kindly submitted
by U.B.Kaplan, 27-AUG-1990.
Lederle Laboratories
60B-317
Pearl River, NY 10965.
location/Qualifiers
1..1843
/organism="Human papillomavirus"
/mol_type="genomic DNA"
/db_xref="taxon:10566"
1023..1499
/gene="E6"
1023..1499
/gene="E6"
/codon_start=1
/protein_id="AAA46973.1"
/db_xref="GI:333060"
/translation="MARFDDPQRPYKLPDLCTELNTSLQVSNACVYCKATLERTEV
YQAPFDLCVYRDCIAYAAACHKCIDPYSIRIETARYSNVSGETLEKINTLVNLL
IRCLRCQKPLNPEKARHLKDKRFRSHINGQYRGQCNCTCDQARQERLRRTQV"

FEATURES
source
gene
CDS

Query Match 95.0%; Score 30.4; DB 10; Length 1843;
Best Local Similarity 96.9%; Pred. No. 0.014;
Matches 31; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TTATTAATAGTGCTGCGGTCCAGAAACC 32
1323 TTGTTAATAAGTGCTGCGGTCCAGAAACC 1354

RESULT 36
HPV45 7858 bp DNA linear VRL 18-Apr-2005
LOCUS Human papillomavirus type 45 genomic DNA.
DEFINITION X74479.1 GI:397022
ACCESSION E1 gene; E2 gene; E4 gene; E6 gene; E7 gene; early protein; L1
VERSION K270003
KEYWORDS
SOURCE Human papillomavirus type 45
ORGANISM Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
Alphapapillomavirus.
REFERENCE 1 (bases 1 to 7858)
AUTHORS Delius, H. and Hofmann, B.
TITLE Primer-directed sequencing of human papillomavirus types
JOURNAL Curr. Top. Microbiol. Immunol. 186, 13-31 (1994)
PUBMED 8205838
REFERENCE 2 (bases 1 to 7858)
AUTHORS Delius, H.
TITLE Direct Submision
JOURNAL Submitted (06-AUG-1993) H. Delius, Deutsches
Krebsforschungszentrum, Abteilung ATV, Im Neuenheimer Feld 506, W
6900 Heidelberg, FRG

FEATURES
source
location/Qualifiers
1..7858
/organism="Human papillomavirus type 45"
/mol_type="genomic DNA"
/db_xref="taxon:10593"
/clone="insert in HindIII site of pGEM-1"
102..578
/gene="E6"
102..578
/gene="E6"
/codon_start=1
/product="early protein"
/protein_id="CA52573.1"
/db_xref="GI:397023"
/db_xref="GOA:P21735"
/db_xref="InterPro:IPR001334"
/translation="MARFDDPKORPYFLPDLCTELNTSLDVSIAACYCAATLERTV
YQAFADLCIVRDCIAVACHCIDIYSRIELRYNSVSGETLEKINTLTYNL
IRLCRCQKPLNPAKRRHMLDKRRFHSIAGYQCCNTCCDQAROEIRRRRETQV"
587..907
/gene="E7"
587..907
/gene="E7"
/codon_start=1
/product="early protein"
/protein_id="CA52574.1"
/db_xref="GI:397024"
/db_xref="GOA:P21736"
/db_xref="InterPro:IPR000148"
/translation="MHGPRETQEIYHLEPQNELDPVDLLCYEQLSSESEENDEADG
VSHAQLEPARAREPQRKILCVCCCKDRIRLTVSSAEDRLTQLFLSTLSFVCPWC
ATNQ"

gene 914..2845
/gene="E1"
914..2845
/gene="E1"
/codon_start=1
/product="early protein"
/protein_id="CA52575.1"

gene 914..2845
/gene="E1"
914..2845
/gene="E1"
/codon_start=1
/product="early protein"
/protein_id="CA52575.1"

/db_xref="GI:397025"
/db_xref="GOA:P36728"
/db_xref="InterPro:IPR001177"
/db_xref="UniProtKB/Swiss-Prot:P36728"
/translation="MADPESTGEGTGNMFVETVETVEKTDVLSDDDETATDRC
SDMVDLIDTQLSICEQAEQTAQALPHADEVQNDADVLIHLKRFPGGSKSENPPLBQ
LSVDTLSPRLQELISLNSGHHKAKARLFTISDGYGSEVEAETQVTVTNBNGS
VHSITQSGGSDSNAENVDPHCSITELKELQSNKKAAALAVFKDIYGLSPFDLYRN
FKSDTKCTDMWVALFGVNPVTVAEGFKTILKPAVLVHIOCLCKNGVILALIRYC
GKRUTVAKGLSTLHVPERCMLIEPRKLASSYALALYRTTGISNISEVSDPEWQ
RLTIIQHGIDDSNFDLSDWQAPNDLIDESMAFQYQALADCNSPAALFLSNCOA
KYLKDAVMCRHYRQAKRQNMNSQWIKYRCSKIDEGDWRPIVQFLRYGVEISFL
RALKEFLKGPFPKNCILLYGPANTGSKYFEGMSFIHFLQGLIISFVNSHFHLEPLAD
TKYAMDADATHTCWTGYEDNVRNALDGNPISIRKHPILLOKCPILITLSNIDPAD
MKPPLRESRVVTFPPHAPPFDKNGNPVEINDKMKCFERTRSLDLHEDBDADAT
EGIPGTFKCYTQNTKPL"
2769..3875
/gene="E2"
2769..3875
/gene="E2"
/codon_start=1
/product="early protein"
/protein_id="CA52576.1"
/db_xref="GI:397026"
/db_xref="GOA:P36794"
/db_xref="InterPro:IPR000427"
/db_xref="InterPro:IPR001866"
/db_xref="InterPro:IPR009021"
/translation="MKMQPKESLSERLSALQKIDHYENDSKDINSQISWQILRL
ENALFTAREHGITKLNHQVPPINISKAHKAIELOMALKGLAQSKYNNIEWTLQD
TCEMLNTPERSOCFFKGGKTVHYVPGNKNQCNVYWDISYITTEGIMDKYACYS
YMGVYIKSDDTTYVQFKSECKEKNSTWRYOYGNVYDNDMSSTDDPTYSARQ
IVRQLDASTSTPKTASVGTPEPHIOTPAIKRQOGLTQHHGRVTVHGNPLTCS
TSNNKRRKVCSGNTTPIILKDGKNSLCLRYRLRKADHYSELSTWHTGCKNKG
ILTVYNSVEQRTPLDVVTIPNSVQISVGWMTI"
4236..5627
/gene="L2"
4236..5627
/gene="L2"
/codon_start=1
/product="late protein"
/protein_id="CA52577.1"
/db_xref="GI:397028"
/db_xref="GOA:P36761"
/db_xref="InterPro:IPR00784"
/translation="MYSRRAARRKASATDLVTRCKQSGTCCPPVINKVEGTLADKI
LQWSIGIFLGIGIGSGSGGRTGYVPLGGSNTVVDGPRRPVIVIEPVGTDS
ITVLVEDSSVYASGAPVPTTSGREITSSTTTPRAVDITTPVDSVSISSSTFNP
AFSDPEITIEPQGEVSAGNLFVETPTSGSHGIEIIPQTPASSGSGTEPLISSTPLPV
RVRVGRGLYSRAQOVRVSTQFLTHRSSLVTFGNDPAVEFLDTLLSEPTSNVPSDF
MDIIRLARPLASSRGVAFVRSRLGQRAIMTRSGKQGGVHFYHDISPIAAVEIEL
OPLISATNDSLDLFDVYADFPFPASTPSTTHKSEFTYKSLTWPSTASSYSNTVPL
TSAMDVPIYTGPDIIILPSHTPMWPSPTNASTTYYIGIHGTQYLLMPWYVYPPKKKK
RIPYPRADGRVA"
5350..7149
/gene="L1"
5350..7149
/gene="L1"
/codon_start=1
/product="late protein"
/protein_id="CA52578.1"
/db_xref="GI:397029"
/db_xref="GOA:P36741"
/db_xref="InterPro:IPR002210"
/db_xref="InterPro:IPR008975"
/db_xref="UniProtKB/Swiss-Prot:P36741"
/translation="MAHNIYGHGIIIFLEKNVVPPIFLQMALMRPSDSTVYLPPEV
ARVSTDDVVSRTSIFYHAGSSRLTLVGNVEYFVVPVPGAGNKQAVPKNVAYQKRFV
ALPDPNKFGLDYSITINPETORLVMAVCVMEIGRGPLGIGLGGHPVYNLDDTESAH
AATAVITQDVNDVSVYKOTQCLICGCVAISEHNAKGLTCKPAQLOPDCPLBLK
NTIIEQDMVDTGYGAMDSEFTLDTKCEVVLIDICQSLCKKPYLQNSADRYGSMFEC

Query Match 95.0%; Score 30.4; DB 5; Length 8039;
 Best Local Similarity 96.9%; Pred. No. 0.012;
 Matches 31; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TTATTAAATAGGTCCTGCGGTGCCAGAAACC 32
 |||||
 Db 6428 TTGTTAAATAGGTCCTGCGGTGCCAGAAACC 6459

RESULT 38
 MUSRP 886 bp mRNA linear ROD 27-APR-1993
 LOCUS Mus musculus ribosomal protein mRNA, complete cds.
 DEFINITION M55235 M38426
 ACCESSION M55235.1 GI:200769
 VERSION ribosomal protein.
 KEYWORDS Mus musculus (house mouse)
 SOURCE Mus musculus
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muridea; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 886)
 AUTHORS Zheng, Y.-M., Su, P., Wang, J.-X., Basu, A., Bhat, K.S., Carter, R. and Avadhani, N.G.
 TITLE Immunochemical and structural similarity between a mouse cytochrome oxidase subunit and ribosomal protein- L7
 JOURNAL Unpublished (1991)
 COMMENT Original source text: Mus musculus (strain Swiss) cDNA to mRNA.
 FEATURES location/Qualifiers
 source 1..886
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="Swiss"
 /db_xref="taxon:10090"
 1..886
 /gene="ribosomal protein"
 11..823
 /gene="ribosomal protein"
 /codon_start=1
 /product="ribosomal protein"
 /protein_id="AA04064.1"
 /db_xref="GI:200770"
 /translation="MEAVPEKKKVAATVPGTLKKKVPAMAKNSKKVPAVETKKKK
 RNPAELVKRLRKKKATKTKVAKRKLIYEAKKHKEVQMTETIRMAKAKN
 PYPAPBKLAFAVIRIRGINGSPPRYKYLQLRKQIPNGTFVAKASIMLAIVDP
 YIANGYPNLKSVNELIKRGYKINKRIATDLSLARSIGKGIICMEDLHEIYT
 VGRKFEKANNFLMPFKLSPRGKKKKTTHFEQGDAGNREDQINLRIRRN"
 886
 /gene="ribosomal protein"

ORIGIN
 polyA_site
 886
 /gene="ribosomal protein"

Query Match 76.9%; Score 24.6; DB 6; Length 886;
 Best Local Similarity 87.1%; Pred. No. 6.7;
 Matches 27; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 TATTATTAAGTGCCTGCGGTGCCAGAAACC 32
 |||||
 Db 97 TAAGAAAAAGTCTCTGCGGTGCCAGAAACC 127

RESULT 39
 BC056907 1225 bp mRNA linear PRI 08-OCT-2003
 LOCUS Homo sapiens cDNA clone MGC:65154 IMAGE:5122136, complete cds.
 DEFINITION BC056907
 ACCESSION BC056907.1 GI:34784913
 VERSION MGC.
 KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homiinae; Homo.
 REFERENCE 1 (bases 1 to 1225)

AUTHORS Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.P., Zeeberg, B., Buettner, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Tothiyuk, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.O., Malek, U.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.U., Hu, Y., S.W., Vallaloo, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S., Sanchez, A., Whitting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butlerfield, Y.S., Krzywinski, M.I., Skalska, U., Smalins, D.E., Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
 Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

1247/932
 2 (bases 1 to 1225)
 Strausberg, R.
 Direct Submission
 Submitted (25-AUG-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: gcgabs-remail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Genome Sequence Center,
 BC Cancer Agency, Vancouver, BC, Canada
info@cgsc.bc.ca

Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butlerfield, Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin, Leticia Hsiao, Martin Krzywinski, Reta Kutische, Oliver Lee, Soo Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven Nees, Pawan Pandoh, Anna-Lisa Prabhu, Parvaneh Saeei, Jacqueline Schein, Duane Smalins, Michael Smith, Lorraine Spence, Jeff Stott, Michael Thorne, Miranda Tsai, Nataja van den Bosch, Jill Vardy, George Yang, Scott Zuyderduyn, Marco Zetter.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>
 Series: IRAC Plate: 119 Row: C Column: 8
 This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis, similarity but not identity to protein.
 location/Qualifiers
 1..1225
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="MGC:65154 IMAGE:5122136"
 /cissue_type="Cervix, Carcinoma"
 /clone_id="NIH MGC_12"
 /lab_host="DH10B"
 /note="Vector: pCMV-SPORT6"
 282..599
 /codon_start=1
 /product="Unknown (protein for MGC:65154)"
 /protein_id="AAH56907.1"
 /db_xref="GI:34784914"
 /translation="MNGPRTADIVLHLEPQNEIPVPLCHROLSPRENDRIDCV
 NQGLPARRAREPQRTHTLWCMCKCKBARIKLVESADDLAFOOLPLNTLTSFVCPWCA
 SQQ"
 282..587
 /note="E7; Region: E7 protein, Early protein"

FEATURES
 source
 CDS
 misc_feature

ORIGIN	/db_xref="CDD:pfam00527"
Query Match	75.6%; Score 24.2; DB 5; Length 1225;
Best Local Similarity	89.7%; Pred. No. 9.8;
Matches	26; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Oy	4 TTAATAAGTGCTGCGGTGCCAGAAACC 32 TTACGAGTGCCTGCGGTGCCAGAAACC 128
Db	100 TTACGAGTGCCTGCGGTGCCAGAAACC 128
RESULT 40	
HUMHELA	1300 bp mRNA linear VRL_06-DEC-1999
LOCUS	Human papillomavirus type 18 proteins E6 and E7 mRNA, complete
DEFINITION	cde.
ACCESSION	M20324
VERSION	M20324.1 GI:183933
KEYWORDS	protein E6; protein E7.
SOURCE	Human papillomavirus type 18
ORGANISM	Human papillomavirus type 18
REFERENCE	Vinuesa, deadNA viruses, no RNA stage; Papillomaviridae; AlphaPapillomavirus. 1 (bases 1 to 1300)
AUTHORS	Inagaki,Y., Tsunokawa,Y., Takebe,N., Nawa,H., Nakaniishi,S., Terada,M. and Sugimura,T.
TITLE	Nucleotide sequences of cDNAs for human papillomavirus type 18 transcripts in HeLa cells
JOURNAL	J. Virol. 62 (5), 1640-1646 (1988)
PUBMED	2833614
COMMENT	The integrated viral sequence starts at position 1.
FEATURES	location/Qualifiers
source	1..724
	/organism="Human papillomavirus type 18"
	/mol_type="mRNA"
	/db_xref="taxon:333761"
	/map="841 bp upstream from Ball site"
	/clone="20"
	/cell_line="HeLa; SKG-[I,II]"
	/focus
	725..1300
	/organism="Homo sapiens"
	/mol_type="mRNA"
	/db_xref="taxon:9606"
	82..255
	/codon_start=1
	/product="protein B6"
	/protein_id="AAA9512.1"
	/db_xref="GI:306834"
	/translation="MARFEDPTRRPPYLPLDLCELNTSLDIEITCYCKTVLELTV PAVEVYSSRKT"
	385..702
	/codon_start=1
	/product="protein E7"
	/protein_id="AAA9513.1"
	/db_xref="GI:306835"
	/translation="MHGPXATLDIVLHLPQNEIPVDLCHEQLSPSEENDEIDGV NHQHLPARRASPORHTMLCMCCKEARIKLVNESSADLRAPQOLPINTLSFYCPWCA SQO"
CDS	
ORIGIN	
Query Match	75.6%; Score 24.2; DB 10; Length 1300;
Best Local Similarity	89.7%; Pred. No. 9.8;
Matches	26; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Oy	4 TTAATAAGTGCTGCGGTGCCAGAAACC 32 TTACGAGTGCCTGCGGTGCCAGAAACC 231
Db	203 TTACGAGTGCCTGCGGTGCCAGAAACC 231
RESULT 41	
AK084350	

LOCUS	AR084350	1095 bp	DNA	linear	PAT 01-SEP-2000
DEFINITION	Sequence 12 from patent US 5981173.				
ACCESSION	AR084350				
VERSION	AR084350.1	GI:10011121			
KEYWORDS	.				
SOURCE	Unknown.				
ORGANISM	Unknown.				
REFERENCE	Unclassified.				
AUTHORS	1 (bases 1 to 1095)				
TITLE	Orch.G., Beaudenon,S. and Longuet,M. Genital human papillomavirus type 68a (HPV-68a), related to the potentially oncogenic HPV-39				
JOURNAL	Patent: US 5981173-A 12 09-NOV-1999; location/Qualifiers				
FEATURES	1..1095				
source	/organism="unknown" /mol_type="unassigned DNA"				
ORIGIN					
Query Match	75.0%;	Score 24;	DB 2;	Length 1095;	
Best Local Similarity	84.4%;	Pred. No. 12;	Mismatches	5;	Gaps 0;
Matches	27;	Conservative	0;		
Oy	1 TTATTATAAGGTGCTCGCGTGCACAACC 32				
Db	580 TTATTGATTAAGTGTCATGAGTTGCCCTAACC 611				
RESULT 42					
LOCUS	AR084344	1108 bp	DNA	linear	PAT 01-SEP-2000
DEFINITION	Sequence 6 from patent US 5981173.				
ACCESSION	AR084344				
VERSION	AR084344.1	GI:10011115			
KEYWORDS	.				
SOURCE	Unknown.				
ORGANISM	Unknown.				
REFERENCE	Unclassified.				
AUTHORS	1 (bases 1 to 1108)				
TITLE	Orch.G., Beaudenon,S. and Longuet,M. Genital human papillomavirus type 68a (HPV-68a), related to the potentially oncogenic HPV-39				
JOURNAL	Patent: US 5981173-A 6 09-NOV-1999; Location/Qualifiers				
FEATURES	1..1108				
source	/organism="unknown" /mol_type="unassigned DNA"				
ORIGIN					
Query Match	75.0%;	Score 24;	DB 2;	Length 1108;	
Best Local Similarity	84.4%;	Pred. No. 12;	Mismatches	5;	Gaps 0;
Matches	27;	Conservative	0;		
Oy	1 TTATTATAAGGTGCTCGCGTGCACAACC 32				
Db	593 TTATTGATTAAGTGTCATGAGTTGCCCTAACC 624				
RESULT 43					
LOCUS	AR084343	1110 bp	DNA	linear	PAT 01-SEP-2000
DEFINITION	Sequence 5 from patent US 5981173.				
ACCESSION	AR084343				
VERSION	AR084343.1	GI:10011114			
KEYWORDS	.				
SOURCE	Unknown.				
ORGANISM	Unknown.				
REFERENCE	Unclassified.				
AUTHORS	1 (bases 1 to 1110)				
TITLE	Orch.G., Beaudenon,S. and Longuet,M. Genital human papillomavirus type 68a (HPV-68a), related to the potentially oncogenic HPV-39				
JOURNAL	Patent: US 5981173-A 5 09-NOV-1999;				

FEATURES
source
Location/Qualifiers
1. 1110
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 75.0%; Score 24; DB 2; Length 1110;
Best Local Similarity 84.4%; Pred. No. 12;
Matches 27; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 TTATTATTAAGTGCTGCGGTGCCAGAAACC 32
594 TTATCATTAAGTGCTGCTGTTGCTGTAACC 625

RESULT 44
AR084345 1110 bp DNA linear PAT 01-SEP-2000
LOCUS
DEFINITION Sequence 7 from patent US 5981173.
ACCESSION AR084345
VERSION AR084345.1 GI:10011116
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.

REFERENCE
1 (bases 1 to 1110)
Orth.G., Beaudenon,S. and Longuet,M.
Genital human papillomavirus type 68a (HPV-68a), related to the
potentially oncogenic HPV-39
Patent: US 5981173-A 7 09-NOV-1999;
Location/Qualifiers
1. 1110
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 75.0%; Score 24; DB 2; Length 1110;
Best Local Similarity 84.4%; Pred. No. 12;
Matches 27; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 TTATTATTAAGTGCTGCGGTGCCAGAAACC 32
596 TTATCATTAAGTGCTGCTGTTGCTGTAACC 627

RESULT 45
HPV66
LOCUS 1313 bp DNA linear VRL 02-JUN-1996
DEFINITION Human Papilloma Virus E6 gene for transforming protein.
ACCESSION X67160
VERSION X67160.1 GI:1197491
KEYWORDS E6 gene; transforming gene.
SOURCE Human papillomavirus type 68
ORGANISM Human papillomavirus type 68
Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
Alphapapillomavirus;
1 (bases 1 to 1313)
Longuet,M., Beaudenon,S. and Orth,G.
Two novel genital human papillomavirus (HPV) types, HPV68 and
HPV70, related to the potentially oncogenic HPV39
J. Clin. Microbiol. 34 (3), 738-744 (1996)

REFERENCE
1 (bases 1 to 1313)
Longuet,M.
Direct Submission
Submitted (03-JUL-1992) M. Longuet, Institut Pasteur, Laboratoire
des Papillomavirus, 25 Rue du Dr Roux, 75024 Paris Cedex 15, FRANCE
Location/Qualifiers
1. 1313
/organism="Human papillomavirus type 68"
/mol_type="genomic DNA"
/db_xref="taxon:45240"
/cell_line="intraepithelial neoplasia"

FEATURES
source

gene 280..1095
/gene="E6"
280..756
CDS
/gene="E6"
/codon_start=1
/product="transforming protein"
/protein_id="CAA47632.1"
/db_xref="GI:1197492"
/db_xref="GOA:P54667"
/db_xref="UniProtKB/Swiss-Prot:P54667"
/translation="MALFNPERPYLPLQCRLLDTLTDVITDVCYCRQLQRTV
YEPASFDLCVVRDGVFPACOSCIKRYAIRIERYSESVATLTETINTKLYNL
IRMSCKPLCPAPKJLHLLTKRLHKLIAGNFTGQCHCWTSKEDRRRIROETQV"
763..1095
/gene="E6"
/codon_start=1
/product="transforming protein"
/protein_id="CAA47633.1"
/db_xref="GI:1197493"
/db_xref="GOA:P54668"
/db_xref="UniProtKB/Swiss-Prot:P54668"
/translation="MGPKPTVDRIYELCPYRIQPVLDVYCHQLDSDDEIDEPDH
AVNHQHLIARDEQRRRIQCLCKCNFALDLVEASRDNLTKLQOLFMDSLNFCV
PWCATETQ"

ORIGIN
Query Match 75.0%; Score 24; DB 10; Length 1313;
Best Local Similarity 84.4%; Pred. No. 12;
Matches 27; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 TTATTATTAAGTGCTGCGGTGCCAGAAACC 32
580 TTATTGATTAAGTGCTGCTGTTGCTGTAACC 611

RESULT 46
AR084352 3283 bp DNA linear PAT 01-SEP-2000
LOCUS
DEFINITION Sequence 14 from patent US 5981173.
ACCESSION AR084352
VERSION AR084352.1 GI:10011123
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.

REFERENCE
1 (bases 1 to 3283)
Orth.G., Beaudenon,S. and Longuet,M.
Genital human papillomavirus type 68a (HPV-68a), related to the
potentially oncogenic HPV-39
Patent: US 5981173-A 14 09-NOV-1999;
Location/Qualifiers
1. 3283
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 75.0%; Score 24; DB 2; Length 3283;
Best Local Similarity 84.4%; Pred. No. 11;
Matches 27; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 TTATTATTAAGTGCTGCGGTGCCAGAAACC 32
2754 TTATCATTAAGTGCTGCTGTTGCTGTAACC 2785

RESULT 47
HPU22461 3283 bp DNA linear VRL 31-MAY-1996
LOCUS
DEFINITION Human papillomavirus type 70 E6 protein (E6), E7 protein (E7), and
L1 protein (L1) genes, complete cds.
ACCESSION U22461
VERSION U22461.1 GI:1345087
KEYWORDS

SOURCE		Human papillomavirus type 70
ORGANISM		Viruses; dsDNA viruses, no RNA stage; Papillomaviridae; Alphapapillomavirus.
REFERENCE		1 (bases 1 to 3283)
AUTHORS		Longuet,M., Beaudenon,S. and Orth,G.
TITLE		Two novel genital human papillomavirus (HPV) types, HPV68 and HPV70, related to the potentially oncogenic HPV39
JOURNAL		J Clin. Microbiol. 34 (3), 738-744 (1996)
PUBMED		8904450
FEATURES		2 (bases 1 to 3283)
AUTHORS		Longuet,M.
TITLE		Direct Submission
JOURNAL		Submitted (13-MAR-1995) Michele Longuet, Laboratoire des Papillomavirus, Unite INSERM 190, Institut Pasteur, 25 rue du Dr Roux, Paris 75015, France
LOCATION/Qualifiers		Location/Qualifiers
source		1..3283
gene		/organism="Human papillomavirus type 70"
CDS		/mol_type="genomic DNA"
		/db_xref="taxon:39457"
		26..1546
		/gene="L1"
		26..1546
		/gene="E1"
		/note="putative major capsid protein"
		/codon_start=1
		/product="L1 protein"
		/protein_id="AAC54879.1"
		/db_xref="GI:1345088"
		/translation="MALRRSSDNTVYLPPEPSAVKVNITDDVTTRTGIVYYAGTSRLTLTVGHVYFRVPNVGGRKQELPKVASAYQVRFSLPDPNKFGLPDSLINPDQRLLVMACIGVIGRGQPLGVGSQGPLNRDLDTENSHFSAVSATOSRDVNSVDYKOTLCIIGRCPVPMHSGMACKAKCTTVOOGDCPRLTELVTALIEODMLDTGYGMDFPTLTGETSKBEPIDIOSVCVKPYDYLOMSADYGDSEMFCLRKBOLEFAHFHNRGGVNGVFTIPSELVIRKDPIBRPGSTHYVSPBSRSSWSQSOLCKNKRYVLAHKQHNNGCMTNQLFITVVDTKSNTFLSACTRTALPAVYSPTKRKEIRHVBEDLOTFPOLCTITLTADMAVITHMPALIDNMNIGVTPPEPSALVDTRYVLOSALAICQKAPTEKDPYDILKFMNVVLDKRFSTEILDQFLPGKRFLLVGARRPRITGPKKRPASAKSSASSHKRRKRVSK"
TATA_signal		2419..2425
gene		/note="putative"
		2454..2930
		/gene="E6"
		2454..2930
		/gene="E6"
		/note="putative transforming protein"
		/codon_start=1
		/product="E6 protein"
		/protein_id="AAC54880.1"
		/db_xref="GI:1345089"
		/translation="MARPPNPAPERPYKLFDICTALDDTTLDITTDICYCTQLQOETVEYEFSLDELFVYRNGBPYAACQKCIKPAAKVRELRHYSNSVYATLTSINTXKYDLSIRCSCKPLCEPAKRLRHVTKRRFHQIASGYTQCCHCWTSNRDBRRRIIRETV"
gene		2939..3268
		/gene="E7"
		2939..3268
		/gene="E7"
		/note="putative transforming protein"
		/codon_start=1
		/product="E7 protein"
		/protein_id="AAC54881.1"
		/db_xref="GI:1345090"
		/translation="MHGRPRTLQEIIVLDLYYNRIQPVLYVCHNQBLESDNETEPDPHVVNHQQQLLARREPQRKHQCMCKCNTTLTLVAESQENRSLDLQEMETLSVFCVWCASGTQ"
ORIGIN		
Query Match		75.0%; Score 24; DB 10; Length 3283;
Best Local Similarity		84.4%; Pred. No. 11;
Matches		27; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
1		TTATTAAATGAAGTCGCGGTGCGGTCGCAAGAACC 32

RESULT 48	DB	2754	TTATCATTAAAGGTGCATGAGTTCCTAAACC	2785
LOCUS	DEFINITION	4618 bp	mRNA	linear
DEFINITION	DEFINITION	Viral-cellular fusion mRNA with Human Papillomavirus type 68 B6 and E7 genes, and Homo sapiens APM-1 gene.		
VERSION	VERSION	Y14591.1	GI:3005941	
KEYWORDS	KEYWORDS	APM-1 gene; B6 gene; E7 gene; fusion.		
SOURCE	SOURCE	Human Papillomavirus type 68		
ORGANISM	ORGANISM	Human Papillomavirus type 68		
REFERENCE	REFERENCE	Alphapapillomavirus; no RNA stage; Papillomaviridae; 1 (bases 1 to 4618)		
AUTHORS	AUTHORS	Reuter, S., Dellijs, H., Kahn, T., Hofmann, B., zur Hausen, H. and Schwartz, B.		
TITLE	TITLE	Characterization of a novel human papillomavirus DNA in the cervical carcinoma cell line ME180		
JOURNAL	JOURNAL	J. Virol. 65 (10), 5564-5568 (1991)		
PUBMED	PUBMED	1716694		
REFERENCE	REFERENCE	2		
AUTHORS	AUTHORS	Reuter, S., Bartelmann, M., Vogt, M., Geisen, C., Napiercki, I., Kahn, T., Dellijs, H., Lichter, P., Wetz, S., Korn, B. and Schwartz, B.		
TITLE	TITLE	APM-1, a novel human gene, identified by aberrant co-transcription with papillomavirus oncogenes in a cervical carcinoma cell line, encodes a B7B/POZ-zinc finger protein with growth inhibitory activity		
JOURNAL	JOURNAL	EMBO J. 17 (1), 215-222 (1998)		
PUBMED	PUBMED	9427755		
REFERENCE	REFERENCE	3 (bases 1 to 4618)		
AUTHORS	AUTHORS	Dellijs, H.		
TITLE	TITLE	Direct Submission		
JOURNAL	JOURNAL	Submitted (13-AUG-1997) H. Dellijs, Deutsches Krebsforschungszentrum, Angewandte Tumoriologie, Im Neuenheimer Feld 242, D-69120 Heidelberg, FRG		
REMARK	REMARK	Revised by author 18-DEC-97		
COMMENT	COMMENT	Related sequence M73258.		
FEATURES	FEATURES	Location/Qualifiers		
source	source	1..839		
		/organism="Human papillomavirus type 68"		
		/mol_type="mRNA"		
		/db_xref="taxon:45240"		
		/focus		
		840..4618		
		/organism="Homo sapiens"		
		/mol_type="mRNA"		
		/db_xref="taxon:9606"		
		/chromosome="18"		
		/map="q21"		
		/cell_line="ME180"		
		/tissue_type="uterine cervix"		
		1..477		
		/gene="B6"		
		1..477		
		/gene="B6"		
		/codon_start=1		
		/product="B6 protein"		
		/protein_id="CAA74931.1"		
		/db_xref="GI:3005942"		
		/db_xref="GI:3005942"		
		/db_xref="GGA:O7KX8"		
		/db_xref="InterPro:IPRO01334"		
		/db_xref="UniProtKB/TrEMBL:O7KX8"		
		/translation="MALFNPERPYGLPDLCTLDLTDVITDCYCRRLQRETVYEFPGDINLVYRGVPLAACOSGCKFYAKIRIAYSESVAITLTINTLQYDLSIRMCCKPLSPAPKLRHLNKKRFRKLAAGFTGQCHCWTSKRERDRRTROFQV"		
		485..817		
		/gene="B7"		
		485..817		
		/gene="B7"		
		/codon_start=1		
		/product="B7 protein"		

/protein_id="CAA74932.1"
/db_xref="GI:3005943"
/db_xref="GOA:073452"
/db_xref="InterPro:IPR000148"
/db_xref="UniProtKB/TREMBL:O73452"
/translation="MHGPKPTVOEIVLELCPNEIEBPVDLVCHQGLSDSDDEIDEPDH
AVNHQQLARDEOQRHTIQCTCKCNMLQLVYASRENLRKLELTPMDSINFLVC
PWCAETERO"
856..2715
/gene="APM-1"
856..2715
/gene="APM-1"
/codon_start=1
/product="APM-1 protein"
/protein_id="CAA74933.1"
/db_xref="GI:3005944"
/db_xref="GOA:073453"
/db_xref="InterPro:IPR000210"
/db_xref="InterPro:IPR007087"
/db_xref="UniProtKB/TREMBL:O73453"
/translation="MANDIDELIGIPPNHSEVLTSLNROBDHGLCDVILVVOEBS
YRHRSVLAACSKYFKLFTAGTLAGOPRYEIDFVQPEALAILPEAYTSLTLTNG
NVNHIILAAAMLEIQCIIVNVCLEIMEGSGDEEDDEDDDEDEDEDEDEDEDE
EEBDDDDDEDFADQENLPQDISCHQSPSTKTHLTERKAYSDTRDFPDSFOGSPGH
LGYLRDPSISLRENLXPKANIPIRRSLSPAPDFPLMGDFGAFQALPEOPMD
SGPLDIVIKRKIKKEEKELPPPPPPPPNDPKMDPDLPGSLGIPATKANDYGAV
LNLPSATHLGLGPPPPPLVEERKLKPKASQCCPCHKYMAGKLPNNHATTHGKRY
MCTICEVRFRODLKTHMRKHTGERPYLCHCAKRVNHYDLKNHRHITGVRYOC
EFGYKSPTRSDHHRHKQSCRMARPRRKPAMWASLLFGGSPADPKAFVWP
PALGEVGHGGAVALCPSPKAPFLAARGLASLOELRQFEETOMKLFGRQOLEA
BRNAGGLALALANVAARPYPLPDPMAAGLAGLPGLAGLNHVASMSANN"

ORIGIN

Query Match 75.0%; Score 24; DB 10; Length 4618;
Best Local Similarity 84.4%; Pred. No. 11;
Matches 27; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 TTATTAATAAGTCGCTGCGTCCGAGAAACC 32
Db 301 TTATCAATAAGTCGATGTGCTGGAACC 332

RESULT 49
HMMHPVME18 6042 bp DNA linear VRL 30-NOV-1999
LOCUS
DEFINITION Human cellular DNA/Human papillomavirus proviral DNA.
ACCESSION M73258.1 GI:184383
VERSION M73258.1
KEYWORDS proviral gene.
SOURCE Human papillomavirus
ORGANISM Human papillomavirus
Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
unclassified Papillomaviridae.
1 (bases 1 to 6042)
Reuter,S., Dellus,H., Kahn,T., Hofmann,B., zur Hausen,H. and
Schwarz,E.
Characterization of a novel human papillomavirus DNA in the
cervical carcinoma cell line ME180
J. Virol. 65 (10), 5564-5568 (1991)
1716694
Location/Qualifiers
1..5993
/organism="Human papillomavirus"
/mol_type="genomic DNA"
/db_xref="taxon:10566"
/locus_tag="HPV18"
5994..6042
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/cell_line="ME180"
/cell_type="carcinoma"
/tissue_type="cervix"

source

source

CDS

227..1636
/codon_start=1
/product="L2"
/protein_id="AAP14009.1"
/db_xref="GI:6478869"
/translation="MVSRRARRKRRASATELYTKQSGTCPPDVINKVEGTTLADKL
LQMTSLGIFLAGIGIGSGTGRGTGPIPLAGKPNYVDVSPAPRVIPVGTES
IYQLVDESVYTSSTGTPVPTGSGTPTSTTTTPAVLITPSGSGVOGSSSTNP
APADPTIIVYQGVSGNVPVSTPTSGTGEIEMQVATGCTGPTISNPIRGV
SRVAGPRTIYKRAHQVRVSNFDPVTHPSGTYTDNPAFEEVDVTLITPEADIDPF
LDIVLRHPALTSRGTGRVSRVSKATMTFRGTQGAOVHYHDIISGLAPADISIL
QPLVAPQSDPMDTLVDIYAPDVTNTVLDTAHNAFTSRSHISVPSLASTSTYA
NTPIPIGTANTPEVNTGPDVLPATSPQLPLTSTPIDTYAITIYGTNYLLPLLF
LKKRKRLPYFPADGIVAL"
1637..3334
/codon_start=1
/product="L1"
/protein_id="AAP14010.1"
/db_xref="GI:6478870"
/translation="MALMRSSDMVYLPSPSVAKVNTDVTYRTGIYYVAGSRLLT
VGRPYRVPMSGGRKODIPKVSAYQVRVPISLPDKRSIPSTLYNPNORLVMAK
VGVEIGRQPLVGLSGHPLYNRLDPTENSFPSSKNKPKQSRDVSVDYKQQLCTIG
CVPLAIGEMAKGKSCRSNVQPDCCPELAVNPPIQGDMLDTGYGAMDSTLQETS
EVLVDICQSYCKTDPDYLQMSADYVGSMPFLCBRLQFAHFNRRGAVGDTPLTELY
IKGTDLRDSSSYVYVAPSPGSMVSDSLFNKPYMLHAKQHNNGICMHNQLFTVY
DTRSRNFTLSTTSSAVNPVINDPKRKEVIRHVEEDLQFIQQLCTITLSTVMSYI
HMANPSALIDMNGVAPPPSASIVDTYRIQSAALITQKQAPAPTKDPPIDGLNPNV
NLKRSSELDQPLGKRLQLQGVRRRPTIGPKRPATVTAISTSNNKRRVSK"
3334..4422
/codon_start=1
/product="E6"
/protein_id="AAP14011.1"
/db_xref="GI:6478871"
/translation="MSIPMALFNNPEERPYKLDLCRTLDLTHDVTDCVYCRQLQ
RTEVEYFAFDGLNVVYVDGVPLAACSCIFVYKIRLRYSVVAATLETITNTL
YDSIRCMCTKLKPLSPAEKLRHNSKRRFKIKAGNFGQCRHCTWSREDRRRRQRT
QV"

CDS

4430..4762
/codon_start=1
/product="E7"
/protein_id="AAP14012.1"
/db_xref="GI:6478872"
/translation="MHGPKPTVOEIVLELCPNEIEBPVDLVCHQGLSDSDDEIDEPDH
AVNHQQLARDEOQRHTIQCTCKCNMLQLVYASRENLRKLELTPMDSINFLVC
PWCAETERO"
4769..5119
/codon_start=1
/product="E1a"
/protein_id="AAP14013.1"
/db_xref="GI:6478873"
/translation="MANCGTDDGTCGNGCFPVQAVLVDKQTDYTSDEEDENATNG
SDMWVPIIDATDIDICIAERTAVLLNMQOQSDAQTALRKRYTDSISSPLAKSP
LDELISWKMQLTRR"
5092..6042
/codon_start=1
/product="E1b"
/protein_id="AAP14014.1"
/db_xref="GI:6478874"
/translation="MEVETNSVTVANTNGADGEDEGENDSIREDCSSVDSADISE
NOPKSPITQKVLQCNKKAMLTFFKVVYGLSFVDLVRFTKSDTTCTDWAAAF
GVNPTIABGFKLTKOYALYTHIOCLDTKGLIILMLIRKCGKNTRTVKGSGTLLH
VPSCKMLQDEPKRSPVAAALYWRKGISNVAOCGPPEIKLITIIHQHIDSVPL
SDMVQAFDEHLELDESIDIATFSYAMLADCNSSNAAPLKSNGQAKTVADCATMCRHYKRA
QKRQMSMPQWIKRCSKCDGGRMDCRILYICLLOAS"

CDS

4430..4762
/codon_start=1
/product="E7"
/protein_id="AAP14012.1"
/db_xref="GI:6478872"
/translation="MHGPKPTVOEIVLELCPNEIEBPVDLVCHQGLSDSDDEIDEPDH
AVNHQQLARDEOQRHTIQCTCKCNMLQLVYASRENLRKLELTPMDSINFLVC
PWCAETERO"
4769..5119
/codon_start=1
/product="E1a"
/protein_id="AAP14013.1"
/db_xref="GI:6478873"
/translation="MANCGTDDGTCGNGCFPVQAVLVDKQTDYTSDEEDENATNG
SDMWVPIIDATDIDICIAERTAVLLNMQOQSDAQTALRKRYTDSISSPLAKSP
LDELISWKMQLTRR"
5092..6042
/codon_start=1
/product="E1b"
/protein_id="AAP14014.1"
/db_xref="GI:6478874"
/translation="MEVETNSVTVANTNGADGEDEGENDSIREDCSSVDSADISE
NOPKSPITQKVLQCNKKAMLTFFKVVYGLSFVDLVRFTKSDTTCTDWAAAF
GVNPTIABGFKLTKOYALYTHIOCLDTKGLIILMLIRKCGKNTRTVKGSGTLLH
VPSCKMLQDEPKRSPVAAALYWRKGISNVAOCGPPEIKLITIIHQHIDSVPL
SDMVQAFDEHLELDESIDIATFSYAMLADCNSSNAAPLKSNGQAKTVADCATMCRHYKRA
QKRQMSMPQWIKRCSKCDGGRMDCRILYICLLOAS"

CDS

4430..4762
/codon_start=1
/product="E7"
/protein_id="AAP14012.1"
/db_xref="GI:6478872"
/translation="MHGPKPTVOEIVLELCPNEIEBPVDLVCHQGLSDSDDEIDEPDH
AVNHQQLARDEOQRHTIQCTCKCNMLQLVYASRENLRKLELTPMDSINFLVC
PWCAETERO"
4769..5119
/codon_start=1
/product="E1a"
/protein_id="AAP14013.1"
/db_xref="GI:6478873"
/translation="MANCGTDDGTCGNGCFPVQAVLVDKQTDYTSDEEDENATNG
SDMWVPIIDATDIDICIAERTAVLLNMQOQSDAQTALRKRYTDSISSPLAKSP
LDELISWKMQLTRR"
5092..6042
/codon_start=1
/product="E1b"
/protein_id="AAP14014.1"
/db_xref="GI:6478874"
/translation="MEVETNSVTVANTNGADGEDEGENDSIREDCSSVDSADISE
NOPKSPITQKVLQCNKKAMLTFFKVVYGLSFVDLVRFTKSDTTCTDWAAAF
GVNPTIABGFKLTKOYALYTHIOCLDTKGLIILMLIRKCGKNTRTVKGSGTLLH
VPSCKMLQDEPKRSPVAAALYWRKGISNVAOCGPPEIKLITIIHQHIDSVPL
SDMVQAFDEHLELDESIDIATFSYAMLADCNSSNAAPLKSNGQAKTVADCATMCRHYKRA
QKRQMSMPQWIKRCSKCDGGRMDCRILYICLLOAS"

CDS

4430..4762
/codon_start=1
/product="E7"
/protein_id="AAP14012.1"
/db_xref="GI:6478872"
/translation="MHGPKPTVOEIVLELCPNEIEBPVDLVCHQGLSDSDDEIDEPDH
AVNHQQLARDEOQRHTIQCTCKCNMLQLVYASRENLRKLELTPMDSINFLVC
PWCAETERO"
4769..5119
/codon_start=1
/product="E1a"
/protein_id="AAP14013.1"
/db_xref="GI:6478873"
/translation="MANCGTDDGTCGNGCFPVQAVLVDKQTDYTSDEEDENATNG
SDMWVPIIDATDIDICIAERTAVLLNMQOQSDAQTALRKRYTDSISSPLAKSP
LDELISWKMQLTRR"
5092..6042
/codon_start=1
/product="E1b"
/protein_id="AAP14014.1"
/db_xref="GI:6478874"
/translation="MEVETNSVTVANTNGADGEDEGENDSIREDCSSVDSADISE
NOPKSPITQKVLQCNKKAMLTFFKVVYGLSFVDLVRFTKSDTTCTDWAAAF
GVNPTIABGFKLTKOYALYTHIOCLDTKGLIILMLIRKCGKNTRTVKGSGTLLH
VPSCKMLQDEPKRSPVAAALYWRKGISNVAOCGPPEIKLITIIHQHIDSVPL
SDMVQAFDEHLELDESIDIATFSYAMLADCNSSNAAPLKSNGQAKTVADCATMCRHYKRA
QKRQMSMPQWIKRCSKCDGGRMDCRILYICLLOAS"

ORIGIN

Query Match 75.0%; Score 24; DB 10; Length 6042;
Best Local Similarity 84.4%; Pred. No. 10;
Matches 27; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 TTATTAATAAGTCGCTGCGTCCGAGAAACC 32
Db 4246 TTATCAATAAGTCGATGTGCTGGAACC 4277

RESULT 50
DQ080079
LOCUS DQ080079 7822 bp DNA circular VRL 06-DEC-2005
DEFINITION Human papillomavirus type 68a, complete genome.
VERSION DQ080079
ACCESSION DQ080079.1 GI:71726685
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
PUBMED
REFERENCES
AUTHORS
TITLES
JOURNAL
FEATURES
source
1. 7822
/organism="Human papillomavirus type 68a"
/mol_type="genomic DNA"
/isolate="Plasmid 952"
/isolation_source="Biopsy of low-grade CIN"
/db_xref="taxon:338322"
/country="France"
/note="type: 68a; acronym: HPV68a"
1. 477
/codon_start=1
/product="E6 protein"
/protein_id="AA239491.1"
/db_xref="GI:71726686"
/translation="MALFPAHNSRPYKLPDLCTLTDTLTDVITDCVYCGROLQRTVE
YERAPDLCAVYDGVPAACGSCIKYATIRLRYTSBSVATTTATNTTGLNLT
IRMSCKPLCPAPKRLHLLTKRLHKLKAGNFTGQCHWTSRDERRLRQRTQV"
484. 816
/codon_start=1
/product="E7 protein"
/protein_id="AA239492.1"
/db_xref="GI:71726687"
/translation="MHGPKPTVQEIYIELCPYNEIQPVDIYCHQIADSDSDIIBDPH
AVNHQHLIARDBOORHRIQCLCCCKNRAALVVEASRDLRTLLQQLFMSDLNFC
PMCATETQ"
823. 2745
/codon_start=1
/product="E1 protein"
/protein_id="AA239493.1"
/db_xref="GI:71726688"
/translation="MANCGSTDEGDCGNGMFVEKAVLDKQTDVTSBEDENATDTG
SPKWDPLDSTHICIAERETAOVLNMQAORDAORVAALNKGDSISSEPLASP
LQELSNVSTQAROPAYTVDPBGYGMETGSEVAALNKGDSISSEPLASP
EDSSVDSADSENQDPSPTTQKVLQCNKKAALTEPKYVIGSPDVIPTPS
DKTCTDVAIIFGNPTIABGFKTLIKQALTYHIOCLTKGILMLIRKCGN
RITVKGKLTLLHVPDSCLLPPKLSPPAALVWRTGISNISEVGGDTPPMIKRT
IIOHIGIDSVFELSDVQMAFDBELTDDSDIAFOYALADCNAAFLKSNCAKY
KDCATGCRHYKRAQKQMTQWIKFRCSCDEGDMRPVQFLRYGLESITFLCAL
KDLKGTGKRNICIVIHGPNTGSGYPCMSLIHFLOGTIIISVNSASHFLPLADAT
AMLDAGTCQMSYEDNMYRAALDGNPISLDRKHRLIQLICPMLITSNTNPNVDRW
PYLHSLITVFKFPAFPFQNRNRPVYITINDKMKCFEKTWCRLDQDEDBDNDGN
TFPTFKVTGENIKTL"
2672. 33784
/codon_start=1
/product="E2 protein"
/protein_id="AA239494.1"
/db_xref="GI:71726689"
/translation="MKETMETLISQRLNVLQKILBHYEODSKCIDIQHIYMNVCRLLE

MAIYARERGMHNDHQVPPVNIISKTAQALIELQMALESLAKTAYAEWTLADT
SNEIMTKKQCKFKKHGVTVVEWYDGNKSNMAYVVMGTIFYKNSIDTCKQGRDY
WGVYVYEQQKTYERFMOADLYGSGKMDVYNNIITCPSPKSTIDGKYSTES
VADLQNTTATHTTAGACCTKRTNSTRSRYPCQCCGITBSEPNVVDVDDLLRS
AGQNRKRSICCGDTPVYHKGDKNGLKCLARILQCHNALYNDISCTMHMIRGSGTN
TGLIVTVSEARDFLETVLKLPSVTVISLGMIL"
<3267. 3551
/note="Lacks traditional start codon"
/codon_start=1
/product="E4 protein"
/protein_id="AA239495.1"
/db_xref="GI:71726690"
/translation="SIVTLCAVPLTEKYPLNLPLTYRTPPRIPPOVPHAKQOTR
RRGSIIPDSABSLSPSPPTCPWTVSISHSVQAKTXDGAVVVTLH."
3830. 4051
/codon_start=1
/product="E5 protein"
/protein_id="AA239496.1"
/db_xref="GI:71726691"
/translation="MLVFLVFWFCVMYICCTVPLLOSVMCVVYMIFVFEVLVRT
TPPEVPAVYLFPILPMMLHSLTLNTIQ"
4098. 5507
/codon_start=1
/product="L2 protein"
/protein_id="AA239497.1"
/db_xref="GI:71726692"
/translation="WVSHRAARKRASATDLYKTCQSGTSPDVINKYEGTLLDKI
LQNTSIGTIFLAGIGIGTSGTGAGAGYIPLGKPNVTVVSPARPVLEPVPTSP
IYOLVDSVITSGTPVPTPTGSGEITSSSTTPAVDIPSSGSOVTSRPNP
ATDPTITIEPVQGVRSVGFVSTPTSGTHGEHEIMQVPAITGTETEPSRPIQV
SRVAGRLYSRAHQVRSNFDVTPHPSFVFDNAFEVDTLTYEPDLAPDF
LDVLRHRLPALSIRGTVRSRGKAPFTRGTQIGQVHYHDSINTPADSLBL
QPLVABEQADPMNDLYIAPDNTVTFDHFNAFTTRSHISVPSLASASTVYT
NTTIPLGATMNPVTGPDVLPVLPSTTQPLFPSTPIDTTFALITVGSNYLLPLLF
LTKRKHLPYFPIDGVAS"
5488. 7005
/codon_start=1
/product="L1 protein"
/protein_id="AA239498.1"
/db_xref="GI:71726693"
/translation="MALWRAQDMVYVLPSPVAKVYNTDVTYRTGMYYAGSRLLT
VGHPIYKVPMSGKRGKGIKRVSYVQVRYRVTLTPDPNKSVPSTIYNPDQOMWAC
VGEIIRGQPLVGLGHPULYNRLDITENSPPSSNKPQDSRNVAVLXNDQKQCLITG
CVPAIGEHNAKGSCKPTVNOQDCPPLSLVNTPIBDGMIDTGVAMDGTIQETKS
EVPADIQGVCKYPDYLQMSADYVGSMEFLAREQLFAHFMNRGAMVGTIPTMY
IKGTDIRETPSVYVAPSPSGMVSQSDQLENKPYMFKHAGHNNIGICMNOGLFTV
DTRRSNTLSTTSDTSTPVAVYNSNFKETVHVEVEYDLOFIQLOLITLITDVMGI
HTNPAIILDMNFGVAPPPSASIVDTYRIQSAALITQCDAPAPVKKDPIDGLNFMNV
DIKEKSSSELDQPLARKFLQAGVRRRPITIGPRKATATATATTSRHKRVSK"

ORIGIN

Query Match 75.0%; Score 24; DB 10; Length 7822;
Best Local Similarity 84.4%; Pred. No. 10;
Matches 27; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Cy 1 TTATTATAAGTGCCTGCGGTCGCAAAAC 32
Db 301 TTATTATAAGTGCATGAGTTCCCGAAACC 332

Search completed: May 24, 2006, 06:51:31
Job time: 1059.62 secs

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 24, 2006, 05:34:58 ; Search time 300.037 Seconds
(without alignments)
650.664 Million cell updates/sec

Title: US-10-601-913-1
Perfect score: 28
Sequence: 1 GACATTATTGTTATGTTGATGAGAC 28

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 5244920 seqs, 3486124231 residues

Total number of hits satisfying chosen parameters: 10489840

Minimum DB seq length: 0
Maximum DB seq length: 2000000000.

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 1000 summaries

Database : N_Geneseq.8:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*
14: geneseqn2005s:*
15: geneseqn2006s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	28	100.0	28	2	AAT72441 Human pap
2	28	100.0	28	2	AAT72413 Human pap
3	28	100.0	28	2	AAT72414 Human pap
4	28	100.0	28	2	AAT72401 Human pap
5	28	100.0	452	14	ADK26909 Human pap
6	28	100.0	477	10	ADP09607 Human pap
7	28	100.0	477	10	ADP09607 Human pap
8	28	100.0	477	10	ADP09607 Human pap
9	28	100.0	519	2	AAT71834 Human pap
10	28	100.0	519	2	AAT71835 Human pap
11	28	100.0	570	2	AA075470 HPV16 B6/
12	28	100.0	712	1	AA091600 Partial n
13	28	100.0	768	13	ADR47004 Human pap
14	28	100.0	768	15	ADR47004 Human pap
15	28	100.0	776	2	AAT14663 B6/E7 reg
16	28	100.0	779	2	AA089756 Probe seq
17	28	100.0	790	2	AA029389 DNA encod
18	28	100.0	801	2	AAT31833 Human pap

19	28	100.0	822	2	AA078792
20	28	100.0	822	2	AA029781 Proc.D1/3
21	28	100.0	822	14	AED52633
22	28	100.0	879	2	AA078795
23	28	100.0	879	2	AA029784
24	28	100.0	879	2	AA029784
25	28	100.0	939	6	AAD35101 Human pap
26	28	100.0	1000	2	AA008627 HPV-16 fr
27	28	100.0	1005	1	AA091784 DNA probe
28	28	100.0	1116	2	AA078793
29	28	100.0	1116	2	AA029782
30	28	100.0	1116	14	AED52637
31	28	100.0	1173	2	AA078797
32	28	100.0	1173	2	AA029786
33	28	100.0	1173	14	AED52645
34	28	100.0	7840	5	AA055127
35	28	100.0	7902	2	AA09847
36	28	100.0	7904	2	AA094723
37	28	100.0	7904	2	AA033881
38	28	100.0	7904	2	AA099946
39	28	100.0	7904	2	AA099952
40	28	100.0	7904	8	AD003127
41	28	100.0	7904	11	AD058335
42	28	100.0	7904	14	AD0245647
43	26.4	94.3	149	6	AD037149
44	26.4	94.3	459	10	AD037149
45	26.4	94.3	477	10	AA054445
46	26.4	94.3	477	14	AA051113
47	26.4	94.3	543	10	AD031984
48	26.4	94.3	747	12	AD044067
49	26.4	94.3	747	12	AD044065
50	26.4	94.3	747	12	AD044061
51	26.4	94.3	747	12	AD044063
52	26.4	94.3	747	12	AD044069
53	26.4	94.3	747	12	AD044071
54	26.4	94.3	921	3	AA009477
55	26.4	94.3	7904	2	AA094724
56	24.8	88.6	456	2	AB076223
57	22	78.6	23	3	AA098882
58	22	78.6	23	12	AD071283
59	22	78.6	23	13	AD064344
60	20.6	73.6	462	9	ACH25313
61	20.4	72.9	94	12	AD080685
62	20.4	72.9	193	6	AB073532
63	20.4	72.9	193	6	AB073531
64	19.8	70.7	29993	10	AD037662
65	19.8	70.7	29993	10	AD037660
66	19.8	70.7	38342	4	AA046745
67	19.8	70.7	38342	6	ABK31506
68	19.6	70.0	3930	6	ABK93132
69	19.6	70.0	4423	4	AB010172
70	19.6	70.0	5240	4	AB012564
71	19.6	70.0	5643	6	ABK93133
72	19.6	70.0	110000	14	AD046976
73	19.4	69.3	160671	14	AED04683
74	19.2	68.6	346	10	AD090013
75	19.2	68.6	346	12	ADH51192
76	19.2	68.6	346	12	AD010853
77	19.2	68.6	348	4	AA055371
78	19.2	68.6	522	6	ABN61793
79	19.2	68.6	558	14	AD059554
80	19.2	68.6	4010	4	AA055369
81	19.2	68.6	4010	10	AD090020
82	19.2	68.6	4010	12	ADH51209
83	19.2	68.6	4010	12	AD010870
84	19.2	68.6	5536	6	AB032179
85	19.2	68.6	6389	6	AB033961
86	19.2	68.6	110000	12	AD079173
87	19	67.9	305	6	AB076959
88	19	67.9	438	4	AA018326
89	19	67.9	472	11	ACN89177
90	19	67.9	604	6	AB035909
91	19	67.9	604	6	AB035908

AA078792	HPV fusio
AA029781	Proc.D1/3
AED52633	Fusion pr
AA078795	HPV fusio
AA029784	CLYTH-B6-
AED52641	Fusion pr
AAD35101	Human pap
AA008627	HPV-16 fr
AA091784	DNA probe
AA078793	HPV fusio
AA029782	Proc.D1/3
AED52637	Fusion pr
AA078797	HPV fusio
AA029786	CLYTH-B6-
AED52645	Fusion pr
AA055127	Nucleotid
AA09847	Human pap
AA094723	Human pap
AA033881	HPV-16 ge
AA099946	Nucleotid
AA099952	Nucleotid
AD003127	HPV B6 se
AD058335	Human pap
AD0245647	Human pap
AD037149	HPV-PTM r
AA054445	HPV16 B6
AD031984	Human pap
AD044067	Nucleotid
AD044065	Nucleotid
AD044061	Nucleotid
AD044063	Nucleotid
AD044069	Nucleotid
AD044071	Nucleotid
AA009477	Human pap
AA094724	Human pap
AB076223	Human pap
AA098882	Human pap
AD071283	Probe #18
AD064344	Human pap
ACH25313	Human adu
AD080685	Human rec
AB073532	HPV-PTM r
AB073531	HPV-PTM r
AD037662	Human che
AD037660	Human che
AA046745	Tumour su
ABK31506	Signal tr
ABK93132	Human pro
AB010172	Drosophil
AB012564	Drosophil
ABK93133	Human pro
AD046976	Human pho
AED04683	Cancer-as
AD090013	Soybean F
ADH51192	Soybean F
AD010853	Soybean d
AA055371	Nucleotid
ABN61793	Human can
AD059554	Human col
AA055369	Nucleotid
AD090020	Soybean F
ADH51209	Soybean F
AD010870	Soybean d
AB032179	Human imm
AB033961	Human imm
AD079173	KLRT2 gen
AB076959	U urealiti
AA018326	Human bre
ACN89177	Breast ca
AB035909	Oligonuc1
AB035908	Oligonuc1

C 822	17	60.7	32236	4	AAK91250	Adg41590 Human res
C 823	17	60.7	32548	4	ABL29482	Ad197364 Human res
C 824	17	60.7	36119	13	ABD33605	Ad168657 Human ova
C 825	17	60.7	45960	11	ACN44372	Ad15029 Human ova
C 826	17	60.7	48674	11	ACN44400	Ad174497 Human ova
C 827	17	60.7	74849	11	ACN44932	Ad168125 Human ova
C 828	17	60.7	90401	12	ADQ97515	Adh46993 Human inf
C 829	17	60.7	91278	13	ABD33289	Adh10266 Human ova
C 830	17	60.7	91278	14	AD213338	Adh18693 Human adu
C 831	17	60.7	91278	14	AD213338	Adh18693 Human adu
C 832	17	60.7	95394	13	ADV78742	Adh94374 Gene #872
C 833	17	60.7	100610	12	ADQ97608	Adh9507 Human ova
C 834	17	60.7	100954	13	ABD33574	Adh15847 Human ova
C 835	17	60.7	106556	12	ABE01466	Adh2506 Human kid
C 836	17	60.7	108057	14	ADQ97858	Adh136368 RNA elev
C 837	17	60.7	110000	6	ABQ74954_5	Adh126368 Rice ablo
C 838	17	60.7	110000	6	ABE55320_2	Adh16071 Probe #47
C 839	17	60.7	110000	6	ABX08336_13	Adh29756 Human liv
C 840	17	60.7	110000	6	ABA90521_06	Adh04687 Human gen
C 841	17	60.7	110000	6	ABQ67196_0	Adh3050 Human cdn
C 842	17	60.7	110000	6	ABQ69245_24	Adh129477 3' end of
C 843	17	60.7	110000	6	ABR03041_27	Adh139734 Human ova
C 844	17	60.7	110000	10	ADF77343_15	Adh68314 Human gen
C 845	17	60.7	110000	10	ACF67367_45	Abq18690 Oligonuc1
C 846	17	60.7	110000	10	ACF65388_02	Abq18691 Oligonuc1
C 847	17	60.7	110000	10	ACF65388_03	Adh13560 Solid tum
C 848	17	60.7	110000	10	ACF62745_2	Adh53527 Human pro
C 849	17	60.7	110000	12	AD23585_13	Adh53527 Human pro
C 850	17	60.7	110000	12	ADN97989_13	Adh13131 Rice ablo
C 851	17	60.7	110000	12	ADN97989_13	Aah12010 Human cdn
C 852	17	60.7	110000	12	ADN97989_13	Aah12010 Human cdn
C 853	17	60.7	110000	12	ADN97989_13	Aah12010 Human cdn
C 854	17	60.7	110000	12	ADN97989_13	Aah12010 Human cdn
C 855	17	60.7	110000	13	ADN97989_13	Aah12010 Human cdn
C 856	17	60.7	110000	13	ADN97989_13	Aah12010 Human cdn
C 857	17	60.7	110000	14	ADN97989_13	Aah12010 Human cdn
C 858	17	60.7	110000	14	ADN97989_13	Aah12010 Human cdn
C 859	17	60.7	110000	14	ADN97989_13	Aah12010 Human cdn
C 860	17	60.7	110000	14	ADN97989_13	Aah12010 Human cdn
C 861	17	60.7	110000	14	ADN97989_13	Aah12010 Human cdn
C 862	17	60.7	110000	14	ADN97989_13	Aah12010 Human cdn
C 863	17	60.7	110000	14	ADN97989_13	Aah12010 Human cdn
C 864	17	60.7	110000	14	ADN97989_13	Aah12010 Human cdn
C 865	17	60.7	110000	14	ADN97989_13	Aah12010 Human cdn
C 866	17	60.7	110000	14	ADN97989_13	Aah12010 Human cdn
C 867	17	60.7	110000	14	ADN97989_13	Aah12010 Human cdn
C 868	17	60.7	110000	14	ADN97989_13	Aah12010 Human cdn
C 869	17	60.7	110000	14	ADN97989_13	Aah12010 Human cdn
C 870	17	60.7	110000	14	ADN97989_13	Aah12010 Human cdn
C 871	17	60.7	110000	14	ADN97989_13	Aah12010 Human cdn
C 872	17	60.7	110000	14	ADN97989_13	Aah12010 Human cdn
C 873	17	60.7	110000	14	ADN97989_13	Aah12010 Human cdn
C 874	17	60.7	110000	14	ADN97989_13	Aah12010 Human cdn
C 875	17	60.7	110000	14	ADN97989_13	Aah12010 Human cdn
C 876	17	60.7	110000	14	ADN97989_13	Aah12010 Human cdn
C 877	17	60.7	110000	14	ADN97989_13	Aah12010 Human cdn
C 878	17	60.7	110000	14	ADN97989_13	Aah12010 Human cdn
C 879	17	60.7	110000	14	ADN97989_13	Aah12010 Human cdn
C 880	17	60.7	110000	14	ADN97989_13	Aah12010 Human cdn
C 881	17	60.7	110000	14	ADN97989_13	Aah12010 Human cdn
C 882	17	60.7	110000	14	ADN97989_13	Aah12010 Human cdn
C 883	17	60.7	110000	14	ADN97989_13	Aah12010 Human cdn
C 884	17	60.7	110000	14	ADN97989_13	Aah12010 Human cdn
C 885	17	60.7	110000	14	ADN97989_13	Aah12010 Human cdn
C 886	17	60.7	110000	14	ADN97989_13	Aah12010 Human cdn
C 887	17	60.7	110000	14	ADN97989_13	Aah12010 Human cdn
C 888	17	60.7	110000	14	ADN97989_13	Aah12010 Human cdn
C 889	17	60.7	110000	14	ADN97989_13	Aah12010 Human cdn
C 890	17	60.7	110000	14	ADN97989_13	Aah12010 Human cdn
C 891	17	60.7	110000	14	ADN97989_13	Aah12010 Human cdn
C 892	17	60.7	110000	14	ADN97989_13	Aah12010 Human cdn
C 893	17	60.7	110000	14	ADN97989_13	Aah12010 Human cdn
C 894	17	60.7	110000	14	ADN97989_13	Aah12010 Human cdn

```

968 16.8 60.0 1668 12 ADQ24928 AdQ24928 Human sof
969 16.8 60.0 1731 13 ADS47748 AdS47748 Bacterial
970 16.8 60.0 1754 13 ADQ81676 AdQ81676 Plant ful
971 16.8 60.0 1775 5 ADL45638 AdL45638 Human ova
972 16.8 60.0 1776 5 ADL63723 AdL63723 Human ova
973 16.8 60.0 1805 13 ADX62815 AdX62815 Plant ful
974 16.8 60.0 1831 12 ADJ39572 AdJ39572 Plant cdn
975 16.8 60.0 1839 14 AEB18185 AEB18185 Gibberell
976 16.8 60.0 1898 4 AAH14078 AAH14078 Human cdn
977 16.8 60.0 1907 8 ACC51086 ACC51086 Human pla
978 16.8 60.0 1955 5 ADL45938 ADL45938 Human ova
979 16.8 60.0 1977 13 ADX50135 ADX50135 Plant ful
980 16.8 60.0 1987 3 AAC77791 AAC77791 Human can
981 16.8 60.0 2000 11 ACL38144 ACL38144 Rice stre
982 16.8 60.0 2000 11 ACL38250 ACL38250 Rice stre
983 16.8 60.0 2061 11 AAA70228 AAA70228 Plaemodiu
984 16.8 60.0 2108 14 ADZ88698 ADZ88698 Breast sp
985 16.8 60.0 2333 2 AAZ52901 AAZ52901 Human pro
986 16.8 60.0 2333 2 AAZ42222 AAZ42222 Human nor
987 16.8 60.0 2404 3 AAF18201 AAF18201 Lung canc
988 16.8 60.0 2459 3 AAC75871 AAC75871 Human ORF
989 16.8 60.0 2487 4 AAH15763 AAH15763 Human cdn
990 16.8 60.0 2487 5 AAH78315 AAH78315 DNA encod
991 16.8 60.0 2496 6 ABK63419 ABK63419 Rat seque
992 16.8 60.0 2496 11 ADW21623 ADW21623 Rat hepat
993 16.8 60.0 2532 13 ADX10785 ADX10785 Plant ful
994 16.8 60.0 2557 4 AAK94586 AAK94586 Human ful
995 16.8 60.0 2557 12 ADL31482 ADL31482 Full leng
996 16.8 60.0 2639 9 ADB23035 ADB23035 Corn meth
997 16.8 60.0 2695 4 ABL02360 ABL02360 Drosophil
998 16.8 60.0 2714 14 ADX07168 ADX07168 Cyclin-de
999 16.8 60.0 2722 6 AAD24414 AAD24414 Human RNA
1000 16.8 60.0 2898 6 ABL90024 ABL90024 Human pol

```

ALIGNMENTS

```

RESULT 1
AAT72441 ID AAT72441 standard; RNA; 28 BP.
XX
XX AC AAT72441;
XX
XX 09-FEB-1998 (first entry)
XX
XX DE Human Papillomavirus Type 16 PCR primer.
XX
XX KW Human Papillomavirus; probe; target region; genital cancer; HPV;
XX KW cervical smear; PCR primer; ss.
XX
XX OS Synthetic.
XX OS Human papillomavirus.
XX
XX PN EP774518-A2.
XX
XX PD 21-MAY-1997.
XX
XX PF 15-NOV-1996; 96BP-00308264.
XX
XX PR 15-NOV-1995; 95US-0006854P.
XX
XX PA (GENP-) GEN-PROBE INC.
XX
XX PI Gordon P, Brentano ST, Carter NM, Hammond PW;
XX
XX DR WPI; 1997-274349/25.
XX
XX PT Probes for detection of Human Papillomavirus Type 16 and Type 18 - can
XX PT distinguish between Type 16 and 18, associated with genital cancers.
XX
XX PS Claim 27; Page 33; 70pp; English.
XX

```

```

CC Novel hybridisation assay probes have been developed comprising an
CC oligonucleotide which will hybridise under selected conditions to Human
CC Papillomavirus (HPV) Type 16 and/or 18 (but not Types 6, 11, 31, 33, 35,
CC 39, 45, 51, 52, or 58) target nucleic acids to form detectable target;
CC probe duplex. The present sequence represents a specifically claimed HPV
CC PCR primer (amplification oligonucleotide). Oligonucleotides are useful
CC to detect HPV Type 16 and/or 18 in samples e.g. cervical smears, body
CC fluid, and distinguish these from other HPV variants. Papillomaviruses
CC are small DNA viruses and HPV Types 16 and 18 are associated with genital
CC cancers. HPV PCR primers can amplify HPV Type 16 and/or 18 nucleic acid
CC in a sample. HPV Type 16 and/or 18 can be detected by adding a probe and
CC detecting probe:target duplex formation; target nucleic acid 18
CC optionally amplified. Type 16 or 18 can be specifically detected by
CC amplifying nucleic acids with at least one specifically claimed PCR
CC primer. For Type 18 detection, a helper probe may be used.
CC Oligonucleotides can distinguish target sequences from phylogenetically
CC close HPV Types 16 and 18 whilst available detection sera react with
CC antigens of all papillomaviruses
XX
XX SQ Sequence 28 BP; 8 A; 2 C; 6 G; 0 T; 12 U; 0 Other;
XX
XX Query Match 100.0%; Score 28; DB 2; Length 28;
XX Best Local Similarity 57.1%; Pred. No. 0.4;
XX Matches 16; Conservative 12; Mismatches 0; Indels 0; Gaps 0;
XX
OY 1 GACATTATTGTTATGTTGATGGAAC 28
Db 1 GACUAUUGUUAUGUUAUGUUAUGAAC 28
XX
XX
XX RESULT 2
XX AAT72413 ID AAT72413 standard; DNA; 28 BP.
XX
XX AC AAT72413;
XX
XX 09-FEB-1998 (first entry)
XX
XX DE Human Papillomavirus Type 16 PCR primer.
XX
XX KW Human Papillomavirus; probe; target region; genital cancer; HPV;
XX KW cervical smear; PCR primer; ss.
XX
XX OS Synthetic.
XX OS Human papillomavirus.
XX
XX PN EP774518-A2.
XX
XX PD 21-MAY-1997.
XX
XX PF 15-NOV-1996; 96BP-00308264.
XX
XX PR 15-NOV-1995; 95US-0006854P.
XX
XX PA (GENP-) GEN-PROBE INC.
XX
XX PI Gordon P, Brentano ST, Carter NM, Hammond PW;
XX
XX DR WPI; 1997-274349/25.
XX
XX PT Probes for detection of Human Papillomavirus Type 16 and Type 18 - can
XX PT distinguish between Type 16 and 18, associated with genital cancers.
XX
XX PS Claim 23; Page 33; 70pp; English.
XX
XX
XX CC Novel hybridisation assay probes have been developed comprising an
XX CC oligonucleotide which will hybridise under selected conditions to Human
XX CC Papillomavirus (HPV) Type 16 and/or 18 (but not Types 6, 11, 31, 33, 35,
XX CC 39, 45, 51, 52, or 58) target nucleic acids to form detectable target;
XX CC probe duplex. The present sequence represents a specifically claimed HPV
XX CC PCR primer (amplification oligonucleotide). Oligonucleotides are useful
XX CC to detect HPV Type 16 and/or 18 in samples e.g. cervical smears, body
XX CC fluid, and distinguish these from other HPV variants. Papillomaviruses
XX CC are small DNA viruses and HPV Types 16 and 18 are associated with genital
XX CC cancers. HPV PCR primers can amplify HPV Type 16 and/or 18 nucleic acid
XX CC in a sample. HPV Type 16 and/or 18 can be detected by adding a probe and
XX CC detecting probe:target duplex formation; target nucleic acid 18
XX CC optionally amplified. Type 16 or 18 can be specifically detected by
XX CC amplifying nucleic acids with at least one specifically claimed PCR
XX CC primer. For Type 18 detection, a helper probe may be used.
XX CC Oligonucleotides can distinguish target sequences from phylogenetically
XX CC close HPV Types 16 and 18 whilst available detection sera react with
XX CC antigens of all papillomaviruses
XX

```

CC are small DNA viruses and HPV Types 16 and 18 are associated with genital
CC cancers. HPV PCR primers can amplify HPV Type 16 and/or 18 nucleic acid
CC in a sample. HPV Type 16 and/or 18 can be detected by adding a probe and
CC detecting probe:target duplex formation; target nucleic acid is
CC optionally amplified. Type 16 or 18 can be specifically detected by
CC amplifying nucleic acids with at least one specifically claimed PCR
CC primer. For Type 18 detection, a helper probe may be used.
CC Oligonucleotides can distinguish target sequences from phylogenetically
CC close HPV Types 16 and 18 whilst available detection sera react with
CC antigens of all papillomaviruses
CC
CC
SQ Sequence 28 BP; 12 A; 6 C; 2 G; 8 T; 0 U; 0 Other;
XX
XX
Query Match 100.0%; Score 28; DB 2; Length 28;
Best Local Similarity 100.0%; Pred. No. 0.4;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 GACATTATTGTTATAGTTGTATGGAAC 28
DB 28 GACATTATTGTTATAGTTGTATGGAAC 1
XX
XX
RESULT 3
AAT72414/C
ID AAT72414 standard; RNA; 28 BP.
XX
XX AAT72414;
XX
XX 09-FEB-1998 (first entry)
XX
XX Human Papillomavirus Type 16 PCR primer.
XX
XX Human Papillomavirus; probe; target region; genital cancer; HPV;
XX cervical smear; PCR primer; ss.
XX
XX Synthetic.
XX Human papillomavirus.
XX
XX BP774518-A2.
XX
XX 21-MAY-1997.
XX
XX 15-NOV-1996; 96EP-00308264.
XX
XX 15-NOV-1996; 96EP-00308264.
XX
XX 15-NOV-1995; 95US-0006854P.
XX
XX (GENP-) GEN-PROBE INC.
XX
XX Gordon P, Brentano ST, Carter NM, Hammond PW;
XX
XX WPI, 1997-274349/25.
XX
XX
PT Probes for detection of Human Papillomavirus Type 16 and Type 18 - can
XX distinguish between Type 16 and 18, associated with genital cancers.
XX
XX
PS Claim 23; Page 34; 70pp; English.
XX
XX Novel hybridisation assay probes have been developed comprising an
CC oligonucleotide which will hybridise under selected conditions to Human
CC Papillomavirus (HPV) Type 16 and/or 18 (but not Types 6, 11, 31, 33, 35,
CC 39, 45, 51, 52, or 58) target nucleic acids to form detectable target;
CC probe duplex. The present sequence represents a specifically claimed HPV
CC PCR primer (amplification oligonucleotide). Oligonucleotides are useful
CC to detect HPV Type 16 and/or 18 in samples e.g. cervical smears, body
CC fluid, and distinguish these from other HPV variants. Papillomaviruses
CC are small DNA viruses and HPV Types 16 and 18 are associated with genital
CC cancers. HPV PCR primers can amplify HPV Type 16 and/or 18 nucleic acid
CC in a sample. HPV Type 16 and/or 18 can be detected by adding a probe and
CC detecting probe:target duplex formation; target nucleic acid is
CC optionally amplified. Type 16 or 18 can be specifically detected by
CC amplifying nucleic acids with at least one specifically claimed PCR
CC primer. For Type 18 detection, a helper probe may be used.
CC Oligonucleotides can distinguish target sequences from phylogenetically

CC close HPV Types 16 and 18 whilst available detection sera react with
CC antigens of all papillomaviruses
CC
CC
SQ Sequence 28 BP; 12 A; 6 C; 2 G; 0 T; 8 U; 0 Other;
XX
XX
Query Match 100.0%; Score 28; DB 2; Length 28;
Best Local Similarity 100.0%; Pred. No. 0.4;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 GACATTATTGTTATAGTTGTATGGAAC 28
DB 28 GACATTATTGTTATAGTTGTATGGAAC 1
XX
XX
RESULT 4
AAT72401
ID AAT72401 standard; DNA; 28 BP.
XX
XX AAT72401;
XX
XX 09-FEB-1998 (first entry)
XX
XX Human Papillomavirus Type 16 probe.
XX
XX Human Papillomavirus; probe; target region; genital cancer; HPV;
XX cervical smear; ss.
XX
XX Synthetic.
XX Human papillomavirus.
XX
XX BP774518-A2.
XX
XX 21-MAY-1997.
XX
XX 15-NOV-1996; 96EP-00308264.
XX
XX 15-NOV-1996; 96EP-00308264.
XX
XX 15-NOV-1995; 95US-0006854P.
XX
XX (GENP-) GEN-PROBE INC.
XX
XX Gordon P, Brentano ST, Carter NM, Hammond PW;
XX
XX WPI, 1997-274349/25.
XX
XX
PT Probes for detection of Human Papillomavirus Type 16 and Type 18 - can
XX distinguish between Type 16 and 18, associated with genital cancers.
XX
XX
PS Claim 22; Page 33; 70pp; English.
XX
XX Novel hybridisation assay probes have been developed comprising an
CC oligonucleotide which will hybridise under selected conditions to Human
CC Papillomavirus (HPV) Type 16 and/or 18 (but not Types 6, 11, 31, 33, 35,
CC 39, 45, 51, 52, or 58) target nucleic acids to form detectable target;
CC probe duplex. The present sequence represents a specifically claimed HPV
CC PCR primer (amplification oligonucleotide). Oligonucleotides are useful
CC to detect HPV Type 16 and/or 18 in samples e.g. cervical smears, body
CC fluid, and distinguish these from other HPV variants. Papillomaviruses
CC are small DNA viruses and HPV Types
CC 16 and 18 are associated with genital cancers. HPV PCR primers can
CC amplify HPV Type 16 and/or 18 nucleic acid in a sample. HPV Type 16
CC and/or 18 can be detected by adding a probe and detecting probe:target
CC duplex formation; target nucleic acid is optionally amplified. Type 16 or
CC 18 can be specifically detected by amplifying nucleic acids with at least
CC one specifically claimed PCR primer. For Type 18 detection, a helper
CC probe may be used. Oligonucleotides can distinguish target sequences from
CC phylogenetically close HPV Types 16 and 18 whilst available detection
CC sera react with antigens of all papillomaviruses
CC
CC
SQ Sequence 28 BP; 8 A; 2 C; 6 G; 12 T; 0 U; 0 Other;
XX
XX
Query Match 100.0%; Score 28; DB 2; Length 28;
Best Local Similarity 100.0%; Pred. No. 0.4;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY      1 GACATTATGTTATAGTTGTATGGAAC 28
      |||
      1 GACATTATGTTATAGTTGTATGGAAC 28

Db

RESULT 5
ADFX26909
ID      ADFX26909 standard; DNA; 452 BP.
XX
AC      ADFX26909;
XX
DT      05-MAY-2005 (first entry)
XX
DE      Human papillomavirus type 16 E6 delta151 mutant DNA.
XX
KW      cell culture; immortalization; gynecology and obstetrics; Andrology;
KW      cancer; neoplasm; mutant; ds.
XX
OS      Human papillomavirus type 16.
OS      Synthetic.
XX
FH      Key      Location/Qualifiers
FT      misc_difference 3..4
FT      /tag= a
FT      /note= "Wild-type bases 4-25 have been removed"
FT      misc_difference 449..450
FT      /tag= b
FT      /note= "Wild-type bases 472-474 have been removed"
XX
PN      JF2005046117-A.
XX
PD      24-FEB-2005.
XX
XX      31-JUN-2003; 2003JP-00283911.
XX
PR      31-JUN-2003; 2003JP-00283911.
XX
PA      (UYKA-) UNIV KANAZAWA TLO YG.
PA      (KOKU-) KOKURITSU GAN CENT SOCHO.
XX
DR      WPI; 2005-175815/19.
XX
PT      Immortalized endometrial glandular epithelial cell strain containing
PT      exogenous immortalizing gene and maintains properties of non-cancer cell
PT      and typical glandular epithelial cells, useful in field of research of
PT      reproduction.
XX
PS      Claim 5; SEQ ID NO 4; 22pp; Japanese.
XX
XX      The invention relates to a novel immortalized endometrial glandular
XX      epithelial cell strain in which an exogenous immortalizing gene is
XX      introduced and the properties of the cell line as non-cancer typical
XX      glandular epithelial cells, are maintained. The cell line of the
XX      invention may be useful for research of reproduction, particularly for
XX      analyzing the adhesion of the fertilized egg during implantation.
XX      analyzing the effect of steroid hormones, analyzing the mechanism of
XX      multistep canceration of the inner glandular epithelial membrane by
XX      genetic engineering of the cell line which is otherwise non-cancerous,
XX      and analyzing the effect of carcinogenic substances on the cell line. The
XX      cell line maintains the properties of normal somatic cells and the
XX      properties of endometrial glandular epithelial cells. The current
XX      sequence is that of the Human papillomavirus type 16 E6 delta151 mutant
XX      DNA of the invention.
XX
SQ      Sequence 452 BP; 156 A; 70 C; 98 G; 128 T; 0 U; 0 Other;
Query Match      100.0%; Score 28; DB 14; Length 452;
Best Local Similarity 100.0%; Pred. No. 0.41;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      1 GACATTATGTTATAGTTGTATGGAAC 28
      |||
      229 GACATTATGTTATAGTTGTATGGAAC 256
Db

```

```

RESULT 6
ADF09607
ID      ADF09607 standard; cDNA; 477 BP.
XX
AC      ADF09607;
XX
DT      12-FEB-2004 (first entry)
XX
DE      Human papillomavirus 16 E6 cDNA SEQ ID NO:108.
XX
KW      human; protein-protein interaction; virucide; cytostatic; vaccine;
KW      human papilloma virus; HPV; cancer; ss; gene.
XX
OS      Human papillomavirus.
XX
PN      WO2003068940-A2.
XX
PD      21-AUG-2003.
XX
XX      14-FEB-2003; 2003WO-US004594.
XX
PR      14-FEB-2002; 2002US-0356911P.
XX
PA      (CURA-) CURAGEN CORP.
PA      (HOFF ) HOFFMANN LA ROCHE INC.
XX
PI      Jackson A, Ooi CE, Lewin DA, Cuthill S;
XX
DR      WPI; 2003-689668/65.
DR      P-PSDB; ADF09515.
XX
PT      New purified complex comprising a first polypeptide and a second
PT      polypeptide, useful for identifying agents for treating/preventing a
PT      condition involving altered level of the complex e.g. human papilloma
PT      virus infection, or cancer.
XX
XX      Example 3; SEQ ID NO 108; 156pp; English.
XX
XX      The invention relates to a novel purified complex comprising a first
XX      polypeptide and a second polypeptide, where the polypeptides comprise
XX      defined amino acid sequences listed in the specification, and where the
XX      first polypeptide binds to the second polypeptide. A complex of the
XX      invention has virucide and cytostatic activity, and may have a use as a
XX      vaccine. The complex is useful for identifying agents for treating or
XX      preventing a conditions involving altered level of the complex, e.g.
XX      human papilloma virus (HPV) infection, or cancer. The compositions,
XX      antibodies, vectors and methods are useful for treating such diseases.
XX      The sequences shown in ADF09584-ADF09697 represent cDNA's of the
XX      invention.
XX
SQ      Sequence 477 BP; 167 A; 76 C; 103 G; 131 T; 0 U; 0 Other;
Query Match      100.0%; Score 28; DB 10; Length 477;
Best Local Similarity 100.0%; Pred. No. 0.41;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      1 GACATTATGTTATAGTTGTATGGAAC 28
      |||
      251 GACATTATGTTATAGTTGTATGGAAC 278
Db

RESULT 7
AAL54430
ID      AAL54430 standard; DNA; 477 BP.
XX
AC      AAL54430;
XX
DT      03-APR-2003 (first entry)
XX
DE      B6 siRNA sequence.
XX

```


FT	/tag= d	
FT	/transl_except= CCG encodes Gln	
FT	121. .123	
FT	/tag= e	
FT	/transl_except= TTG encodes Lys	
FT	124. .126	
FT	/tag= f	
FT	/transl_except= TGT encodes Pro	
FT	127. .129	
FT	/tag= g	
FT	/transl_except= GAT encodes Tyr	
FT	130. .132	
FT	/tag= h	
FT	/transl_except= TTG encodes Asn	
FT	133. .135	
FT	/tag= i	
FT	/transl_except= TTA encodes Lys	
FT	136. .138	
FT	/tag= j	
FT	/transl_except= ATT encodes Pro	
FT	139. .141	
FT	/tag= k	
FT	/transl_except= ACG encodes Leu	
FT	142. .144	
FT	/tag= l	
FT	/transl_except= TGT encodes Cys	
FT	145. .147	
FT	/tag= m	
FT	/transl_except= ATT encodes Asp	
FT	148. .150	
FT	/tag= n	
FT	/transl_except= AAC encodes Leu	
FT	151. .153	
FT	/tag= o	
FT	/transl_except= TGT encodes Leu	
FT	154. .156	
FT	/tag= p	
FT	/transl_except= CAA encodes Ile	
FT	157. .159	
FT	/tag= q	
FT	/transl_except= AAG encodes Arg	
FT	160. .162	
FT	/tag= r	
FT	/transl_except= CCA encodes Cys	
XX		
PN	MO9619496-A1.	
XX		
XX	27-JUN-1996.	
PD		
XX		
PF	20-DEC-1995; 95MO-AU0000868.	
XX		
PR	20-DEC-1994; 94AU-00000157.	
XX		
XX	(CSLC-) CSL LTD.	
PA	(UYOU) UNIV QUEENSLAND.	
XX		
PI	Edwards SJ, Cox J, Webb EA, Frazer I;	
XX		
DR	WPI, 1996-309518/31.	
XX		
DR	P-PSDB; AAR97562.	
XX		
PT	Vaccine variants of human papilloma virus antigens - contain variants of	
PT	B6 and/or E7 protein, pref. deletion mutants, and are used to treat or	
PT	prevent HPV infection.	
XX		
PS	Example 3; Page 17; 37P; English.	
XX		
CC	A variant of the human papilloma virus (HPV) B6 or E7 protein which	
CC	elicits a humoral and/or cellular immune response against HPV can be used	
CC	in vaccines against HPV or to treat HPV infection. The variant is	
CC	preferably a deletion mutant comprising at least half, and preferably two	
CC	-thirds of full length B6 or E7 protein starting from the N- or C-	
CC	terminal, or is a full length B6 moiety fused to a full length E7 moiety.	
CC		

```

CC The variant optionally has a linkage moiety and a foreign protein or
CC peptide which facilitates the purification of, and enhances the
CC immunogenicity of, the fusion protein. This sequence encodes a fusion
CC between the C-terminal end of B6 and the N-terminal end of E7. The
CC protein is also a deletion mutant generated from the sequence described
CC in AATJ1833. (Updated on 27-AUG-2003 to correct OS field.)
CC
XX
XX Sequence 519 BP, 175 A, 92 C, 113 G, 139 T, 0 U, 0 Other;
SQ
Query Match 100.0%; Score 28; DB 2; Length 519;
Best Local Similarity 100.0%; Pident. No. 0.41;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 1 GACATTAATGTTATAGTTTGTATGGAAC 28
Db 65 GACATTAATGTTATAGTTTGTATGGAAC 92
RESULT 10
AATJ1835
ID AATJ1835 standard; DNA; 519 BP.
XX
XX AATJ1835;
AC
XX 27-AUG-2003 (revised)
DT 11-JAN-1997 (first entry)
XX
XX Human papilloma virus B6/E7 protein variant.
DE Human papilloma virus; E6; E7; deletion mutant; HPV; immune response;
XX humoral immune response; cellular immune response; vaccine; ss.
XX
XX Human papillomavirus.
OS
XX
XX Key Location/Qualifiers
FH 1. 519
FT CDS /*tag= a
FT /product= "B6/E7 fusion protein."
FT
XX
XX WO9619496-A1.
XX
XX 27-JUN-1996.
XX
XX 20-DEC-1995; 95MO-AU000868.
XX
XX 20-DEC-1994; 94AU-00000157.
XX
XX (CSLC-) CSL LTD.
XX (UYOU) UNIV QUEENSLAND.
XX
XX Edwards SJ, Cox J, Webb EA, Frazer I;
XX
XX MPI; 1996-309518/31.
XX P-PsDB; AAR97563.
XX
XX Vaccine variants of human papilloma virus antigens - contain variants of
XX B6 and/or E7 protein, pref. deletion mutants, and are used to treat or
XX prevent HPV infection.
XX
XX Example 3; Page 18; 37pp; English.
XX
XX A variant of the human papilloma virus (HPV) E6 or E7 protein which
XX elicits a humoral and/or cellular immune response against HPV can be used
XX in vaccines against HPV or to treat HPV infection. The variant is
XX preferably a deletion mutant comprising at least half, and preferably two
XX thirds of full length E6 or E7 protein starting from the N- or C-
XX terminal, or is a full length E6 moiety fused to a full length E7 moiety.
XX The variant optionally has a linkage moiety and a foreign protein or
XX peptide which facilitates the purification of, and enhances the
XX immunogenicity of, the fusion protein. This sequence encodes a fusion
XX between the C-terminal end of E7 and the N-terminal end of E6. (Updated
XX on 27-AUG-2003 to correct OS field.)
XX

```

Sequence 519 BP; 167 A; 97 C; 111 G; 144 T; 0 U; 0 Other;

Query Match 100.0%; Score 28; DB 2; Length 519;

Best Local Similarity 100.0%; Pred. No. 0.41;

Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACATTATTGTTATAGTTGTATGAAC 28
DB 458 GACATTATTGTTATAGTTGTATGAAC 485

RESULT 11
AAQ75470
ID AAQ75470 standard; DNA; 570 BP.

AC AAQ75470;

DT 16-OCT-2003 (revised)

DT 25-MAR-2003 (revised)

DT 28-JUN-1995 (first entry)

DE HPV16 B6/E7 encoding region.

KW HPV; HPV16; E6 protein; E7 protein; diagnosis; cervical dysplasia;

KW cervix cancer; ds.

XX Human papillomavirus; strain 16.

XX Key Location/Qualifiers

FT CDS 2..478

FT /tag= a

FT /label= B6_encoding_region

FT 481..570

FT /tag= b

FT /label= E7_encoding_region

PN WO9426934-A2.

PN 24-NOV-1994.

XX 06-MAY-1994; 94WO-US005085.

XX 06-MAY-1993; 93US-00058920.

XX (BAXT) BAXTER DIAGNOSTICS INC.

XX Brown UT;

XX WPI; 1995-006821/01.

XX P-PSDB; AAR63865.

XX Human papilloma virus detection assay - by amplification using self

XX sustained sequence replication and hybridisation with a detector probe.

XX Disclosure; Page 24-26; 79pp; English.

XX The sequences of the B6 and E7 polypeptide-encoding regions of human

XX papillomavirus (HPV) 16 and 18 are given in AAQ75470-71 and the encoded

XX proteins in AAR63865-66, respectively. Probes and primers based on these

XX sequences were used for HPV infection diagnosis; expression of B6 and E7

XX 18 diagnostic for cervical cancer or pre-malignant states. (Updated on

XX 25-MAR-2003 to correct PW field.) (Updated on 16-OCT-2003 to standardise

XX OS field)

XX Sequence 570 BP; 199 A; 93 C; 119 G; 159 T; 0 U; 0 Other;

XX Query Match 100.0%; Score 28; DB 2; Length 570;

XX Best Local Similarity 100.0%; Pred. No. 0.41;

XX Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACATTATTGTTATAGTTGTATGAAC 28
DB 252 GACATTATTGTTATAGTTGTATGAAC 279

RESULT 12

AAAN91600
ID AAAN91600 standard; DNA; 712 BP.

XX AAAN91600;

XX 27-AUG-2003 (revised)

XX 17-JUL-1990 (first entry)

XX Partial nucleotide sequence (5' end) of human papilloma virus (HPV) type

XX 16 (HPV-16).

XX Human papilloma virus; type 16; in situ hybridisation assay;

XX cellular smear; benign cervical wart; cervical cancer.

XX Human papillomavirus.

XX WO8902934-A.

XX 06-APR-1989.

XX 30-SEP-1988; 88WO-US003367.

XX 02-OCT-1987; 87US-00103979.

XX (MICR-) MICROPROBE CORP.

XX Schwartz DE, Adams TH;

XX WPI; 1989-114406/15.

XX Hybridisation test for human papilloma virus in cell smears - by reaction

XX with long labelled probe specific for particular virus types, esp. for

XX examining cervical smears.

XX Disclosure; Page 7; 39pp; English.

XX The patent is for a rapid in situ hybridisation assay for detecting and

XX typing human papilloma virus (HPV) in non-frozen cellular smears fixed to

XX a support in absence of aldehyde-based crosslinking reagents. The assay

XX comprises: (1) combining nucleic acid in the sample with at least one

XX detectable probe able to hybridise with 1 or more HPV types; and (2)

XX are used, eg one for HPV types 6 and 11, associated with benign warts,

XX and one for types 16, 18, 31, 33 and 35, associated with cervical cancer.

XX The assay can differentiate between HPV types. It is esp. used as a

XX secondary test. The probes can be synthesised or cloned. (Updated on 27-

XX AUG-2003 to correct OS field.)

XX Sequence 712 BP; 253 A; 120 C; 155 G; 184 T; 0 U; 0 Other;

XX Query Match 100.0%; Score 28; DB 1; Length 712;

XX Best Local Similarity 100.0%; Pred. No. 0.41;

XX Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACATTATTGTTATAGTTGTATGAAC 28
DB 333 GACATTATTGTTATAGTTGTATGAAC 360

RESULT 13
ADRA47004
ID ADRA47004 standard; DNA; 768 BP.

XX ADRA47004;

XX 18-NOV-2004 (first entry)

XX Human papillomavirus type 16 E7 gene for Dengue virus vaccine.

XX ds; gene; cytosolic; virucide; dengue virus; recombinant replicon;

KM	deletion prem protein; C protein; NS1 protein signal; vaccine;
KM	cervical cancer; viral disease; antigen; dendritic cell; immune response;
KM	human papillomavirus.
XX	
XX	Human papillomavirus type 16.
XX	
XX	Key Location/Qualifiers
FT	1..768
FT	CDS
FT	/tag= a
FT	/product= "HPV-16 E7 protein"
PX	
PX	MOZ004072274-A1.
PD	
PD	26-AUG-2004.
PX	
PF	30-JAN-2004; 2004WO-CN000088.
PX	
PR	30-JAN-2003; 2003CN-00115272.
PR	30-JAN-2003; 2003CN-00115273.
PX	
PA	(SHAN-) SHANGHAI TENGGEN BIOMEDICAL CO LTD.
PA	(TENG-) TENGGEN BIOMEDICAL CO.
PA	(BEIJ-) BEIJING ORIENTAL TENGGEN TECHNOLOGY DEV C.
PX	
PI	Pang X;
DR	WPI: 2004-625870/60.
DR	P-PsDB; ADR47005.
DR	GEMBANK; AP486352.
PT	
PT	Virus-like particle vaccines containing dengue virus recombinant replicon as core for carrier, applicable in preventives or/and remedies for tumors like cervical cancer and viral diseases.
XX	
PS	Example 2; SEQ ID NO 1; 38pp; Chinese.
CC	A dengue virus recombinant replicon has a deletion of the complete coding sequence for prem protein of dengue virus and also includes elements of e.g. the non-coding region in the whole of the 5'-end, the coding region of the front 20 amino acids in the C protein, and the coding region of NS1 protein signal; coding regions of all non-structural proteins. The obtained vaccines are useful in producing preventives or/and remedies for cancer like cervical cancer and viral diseases. Such vaccines can efficiently express antigen in infected cells, which is because dengue virus can infect dendritic cells, and can effectively present antigen to provide immunity effect. Different types of dengue virus can be used to repeatedly produce efficient immune response thereby strengthening the body's immune system against the pathogen that contains such antigen.
CC	Human papillomavirus (HPV) vaccines were prepared by using a gene-expressing system using of the full-length dengue virus cDNA clone (pRS/FLD2). The recombinant virus vectors were transfected into baby hamster kidney (BHK) cells to enable the screening of BHK-21 Tet-off cell lines. This sequence corresponds to the HPV type 16 E7 gene used as the gene of interest in the recombinant replicon of the invention.
SQ	Sequence 768 BP; 255 A; 139 C; 172 G; 202 T; 0 U; 0 Other;
Query Match	100.0%; Score 28; DB 13; Length 768;
Best Local Similarity	100.0%; Pred. No. 0.41;
Matches	28; Conservative 0; Mismatches 0; Indels 0; Gaps 0
OY	1 GACATTATTTGTTATAGTTGTGAAC 28
Db	545 GACATTATTTGTTATAGTTGTGAAC 572
RESULT 14	
AEFA0156	
ID	AEFA0156 standard; DNA; 768 BP.
XX	
AC	AEFA0156;
XX	
DT	23-MAR-2006 (first entry)

XX		Human papillomavirus 16 (HPV-16) E7-E6 oncogene.
DE		Vaccine; virus-like particle; replicon; therapeutic; cancer; cytostatic;
KM		neoplasm; viral infection; virucide; infection; oncogene;
KW		coding sequence; ds.
XX		
OS		Human papillomavirus type 16.
FH		
FT	Key	Location/Qualifiers
FT	CDS	1..768
FT		/tag= A
FT		/partial
FT		/product= "Human papillomavirus 16 (HPV-16) E7-E6
FT		oncoprotein"
FT		/note= "No stop codon"
PN		US2006018928-A1.
PD		
PP		26-JAN-2006.
PF		29-JUL-2005; 2005US-00192923.
PR		30-JAN-2003; 2003CN-00115272.
PR		30-JAN-2003; 2003CN-00115273.
PR		30-JAN-2004; 2004WO-00072274.
PA		(PANG/) PANG X.
PI		Pang X;
PS		WPJ; 2006-109169/11.
DR		P-PDB; AF40157.
DR		GENBANK; AF486352, AF469197, AF472508.
XX		
PT		New recombinant DEN replicons with a deletion of prem, useful for
PT		producing a drug for the prophylaxis and treatment of cancer or viral
PT		infection.
XX		
PS		Example 2; SEQ ID NO 1; 24pp; English.
XX		
CC		The present invention provides a virus-like particle (VLP) vaccine which
CC		contains dengue virus (DEN) recombinant replicon as its core. The DEN
CC		replicon contains exogenous nucleotide sequences such as human
CC		papillomavirus (HPV) antigen proteins, immune regulators or combination
CC		of HPV antigen and immune regulators. The invention is useful for
CC		producing a drug for the prophylaxis and treatment of cancer or viral
CC		infection. The present sequence is a human papillomavirus oncogene.
XX		
SQ		Sequence 768 BP; 255 A; 139 C; 172 G; 202 T; 0 U; 0 Other;
	Query Match	100.0%; Score 28; DB 15; Length 768;
	Best Local Similarity	100.0%; Pred. No. 0.41;
	Matches 28; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
OY		1 GACATTATGTTATAGTTGTATGAAC 28
DB		545 GACATTATGTTATAGTTGTATGAAC 572
RESULT 15		
AATL4663		
ID		AATL4663 standard; DNA; 776 BP.
XX		
AC		AATL4663;
XX		
DT		27-AUG-2003 (revised)
DT		25-MAR-2003 (revised)
DT		10-OCT-1996 (first entry)
XX		
DE		E6/E7 region of Human Papilloma Virus 16 (HPV 16).
KW		Human papilloma virus; HPV; detection; cervical cancer; amplification;

Key	Location/Qualifiers
primer_bind	1..24 /tag= a /note= "Primer BB13 binding site."
misc_binding	30..55 /tag= b /note= "Primer H16-58 binding site."
primer_bind	37..57 /tag= c /note= "Primer BB4 binding site."
primer_bind	454..474 /tag= d /note= "Primer BB114 binding site."
primer_bind	480..503 /tag= e /note= "Primer BB11 binding site."
primer_bind	591..621 /tag= f /note= "Primer BB109 binding site."
primer_bind	658..681 /tag= g /note= "Primer H16-686 binding site."
primer_bind	660..683 /tag= h /note= "Primer BB112 binding site."
primer_bind	715..738 /tag= i /note= "Primer H16-743 binding site."
primer_bind	745..768 /tag= j /note= "Primer H16-773 binding site."
US5506105-A.	
09-APR-1996.	
22-MAR-1994;	94US-00216233.
10-DEC-1991;	91US-00808456.
(DADE-) DADE INT INC.	
Haydock PV;	
WPI, 1996-200273/20.	
Detection of low copy number intracellular markers - by 3SR amplification of target RNA in fixed cells then hybridisation with labelled probe, for detecting human papilloma virus in cervical cells.	
Example 2; Fig 3; 21pp; English.	
An in situ hybridisation assay for detecting an intracellular marker of low copy number in cells comprises fixing the cells to a support using paraformaldehyde; treating the cells with a protease to permeabilise them without altering morphology; adding amplification reagents; incubating the cells at below fifty degrees celsius to perform amplification by self-sustained sequence replication; adding a labelled probe complementary to the region between the primers; washing cells to remove unhybridised probe and then detecting the labelled probe. The E6/E7 region of human papilloma virus (HPV) 16 is used especially to detect mRNA being transcribed from this region which becomes active in transformed cells. The method can be used for the early diagnosis of cervical cancer. Primers used to amplify fragments of the B6/B7 region are described in AAT4664-rt4674. (Updated on 25-MAR-2003 to correct PF field.) (Updated on 27-AUG-2003 to correct OS field.)	
Sequence 776 BP; 260 A; 139 C; 169 G; 208 T; 0 U; 0 Other;	

Query Match	100.0%;	Score 28;	DB 2;	Length 776;
Best Local Similarity	100.0%;	Pred. No. 0.41;		
Matches	28;	Conservative 0;	Mismatches 0;	Indels 0;
Qy	1	GACATTATTGTTATAGTTTGATGGAAC	28	
Db	251	GACATTATTGTTATAGTTTGATGGAAC	278	
RESULT 16				
ID	AAx89756	standard; DNA; 779 BP.		
AC	AAx89756;			
XX				
XX	27-AUG-2003 (revised)			
DT	12-OCT-1999 (first entry)			
XX				
XX	Probe sequence for HPV 16 E6/E7.			
XX				
KM	human papilloma virus; infection; gene expression; probe; detection;			
KW	assay; cancer; virus; HPV; ss.			
XX				
XX	Synthetic.			
OS	Human papillomavirus.			
PN	W09929890-A2.			
XX				
PD	17-JUN-1999.			
XX				
PF	11-DEC-1998; 98WO-US026447.			
XX				
PR	12-DEC-1997; 97US-0069426P.			
PR	05-JAN-1998; 98US-0070466P.			
PR	17-APR-1998; 98US-0082167P.			
XX				
FA	(DIGE-) DIGENE CORP.			
XX				
PI	Lorincz AT;			
XX				
DR	WPI, 1999-443850/37.			
XX				
PT	New method for assessing Human Papilloma Virus (HPV) infection by			
XX	comparison of gene expression levels.			
PS	Disclosure; Fig 5; 35pp; English.			
XX				
CC	This nucleotide probe is specific for the HPV16 Human Papilloma Virus			
CC	(HPV) gene. The degree of HPV infection can be assessed, by measuring the			
CC	levels of expression of genes involved in the diseased state, and			
CC	comparing the expression to each other or to reference genes. (Updated on			
CC	27-AUG-2003 to correct OS field.)			
XX				
SQ	Sequence 779 BP; 259 A; 139 C; 170 G; 211 T; 0 U; 0 Other;			
Query Match	100.0%;	Score 28;	DB 2;	Length 779;
Best Local Similarity	100.0%;	Pred. No. 0.41;		
Matches	28;	Conservative 0;	Mismatches 0;	Indels 0;
Qy	1	GACATTATTGTTATAGTTTGATGGAAC	28	
Db	258	GACATTATTGTTATAGTTTGATGGAAC	285	
RESULT 17				
ID	AAQ29389	standard; DNA; 790 BP.		
XX	AAQ29389;			
AC	AAQ29389;			
XX				
DT	24-OCT-2003 (revised)			
DT	25-MAR-2003 (revised)			
DT	09-MAR-1993 (first entry)			

```

XX DE DNA encoding HPV 16 E6/E7 proteins obtd. by PCR.
XX KM Virus vector; vaccinia virus; papillomavirus; HPV; human; amplification;
XX KM immunotherapeutic; ss.
XX OS Human papillomavirus type 16.
XX FH Key Location/Qualifiers
XX FT CDS 1..789
XX FT /tag= a
XX FT /note= "reading frame 1"
XX FT 2..790
XX FT /tag= b
XX FT /note= "reading frame 2"
XX FT 3..788
XX FT /tag= d
XX FT /note= "reading frame 3"
XX FT 6..482
XX FT /tag= e
XX FT /note= "third reading frame encoding HPV 16 E6"
XX FT 485..781
XX FT /tag= c
XX FT /note= "second reading frame encoding HPV 16 E7"
XX PN W0921636-A1.
XX PD 01-OCT-1992.
XX XX 10-MAR-1992; 92WO-GB000424.
XX PF 14-MAR-1991; 91GB-00005383.
XX PR (IMMU ) IMMUNOLOGY LTD.
XX PA Bourne11 MEG, Inglis SC, Munro AJ;
XX DR WPI, 1992-349219/42.
XX DR P-PSDB; AAR27723, AAR27724, AAR27725.
XX PT Recombinant virus vectors encoding human papillomavirus proteins - for
XX PT treating and vaccinating against HPV infections and conditions caused by
XX PT them, such as cervical cancer.
XX PS Disclosure; Fig 1a; 83pp; English.
XX XX The fragment of DNA contg. the HPV-16 E6/E7 coding region was prepd. by
XX CC PCR from plasmid pBR322/HPV16 (Durst et al., PNAS, 80: 3812 (1983)) using
XX CC oligonucleotides S05 and S06. The DNA prod. has a site for NcoI at the
XX CC beginning of the E6 gene and a SmaI site immediately downstream of the
XX CC termination codon for E7. The E6 and E7 ORFs are fused together to form a
XX CC single continuous ORF via site directed mutagenesis and the immortalising
XX CC potential of E7 is removed by altering two key codons of the HPV E7
XX CC sequence. The single ORF of HPV-16 E6/E7 may be inserted into vaccinia
XX CC virus DNA at neutral sites (pref. by inserting two sets of the DNA in
XX CC opposite orientations to overcome the problem of intertypic
XX CC recombination) to make a recombinant virus vector for use
XX CC immunotherapeutically to activate cells of the immune system against HPV.
XX CC See also AAQ23185-400 and AAQ29450-69. (Updated on 25-MAR-2003 to correct
XX CC PN field.) (Updated on 24-OCT-2003 to standardise OS field)
XX SQ Sequence 790 BP; 262 A; 144 C; 175 G; 209 T; 0 U; 0 Other;

Query Match 100.0%; Score 28; DB 2; Length 790;
Best Local Similarity 100.0%; Pred. No. 0.41;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACATTATGTTAGTTGTATGGAAC 28
DB 256 GACATTATGTTAGTTGTATGGAAC 283

```

```

AAT31833
ID AAT31833 standard; DNA; 801 BP.
XX AC AAT31833;
XX XX 27-AUG-2003 (revised)
XX DT 10-JAN-1997 (first entry)
XX XX
XX DE Human papilloma virus E6/E7 protein variant.
XX KM Human papilloma virus; E6; E7; deletion mutant; HPV; immune response;
XX KM humoral immune response; cellular immune response; vaccine; ss.
XX OS Human papillomavirus.
XX FH Key Location/Qualifiers
XX FT CDS 1..801
XX FT /tag= a
XX FT /product= "E6/E7 fusion protein."
XX PN W09619496-A1.
XX PD 27-JUN-1996.
XX XX 20-DEC-1995; 95WO-AU000868.
XX PF 20-DEC-1994; 94AU-00000157.
XX PR (CSLC-) CSL LTD.
XX PA (UYQU ) UNIV QUEBENS LAND.
XX PI Edwards SJ, Cox J, Webb EA, Frazer I;
XX DR WPI, 1996-309518/31.
XX DR P-PSDB; AAR97561.
XX PT Vaccine variants of human papilloma virus antigens - contain variants of
XX PT E6 and/or E7 protein, pref. deletion mutants, and are used to treat or
XX PT prevent HPV infection.
XX PS Example 1; Page 15-16; 37pp; English.
XX XX A variant of the human papilloma virus (HPV) E6 or E7 protein which
XX CC elicits a humoral and/or cellular immune response against HPV can be used
XX CC in vaccines against HPV or to treat HPV infection. The variant is
XX CC preferably a deletion mutant comprising at least half, and preferably two
XX CC -thirds of full length E6 or E7 protein starting from the N- or C-
XX CC terminal, or is a full length E6 moiety fused to a full length E7 moiety.
XX CC The variant optionally has a linkage moiety and a foreign protein or
XX CC peptide which facilitates the purification of, and enhances the
XX CC immunogenicity of, the fusion protein. This sequence encodes a full
XX CC length E6/E7 fusion protein. (Updated on 27-AUG-2003 to correct OS
XX CC field.)
XX SQ Sequence 801 BP; 266 A; 150 C; 174 G; 211 T; 0 U; 0 Other;

Query Match 100.0%; Score 28; DB 2; Length 801;
Best Local Similarity 100.0%; Pred. No. 0.41;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACATTATGTTATAGTTGTATGGAAC 28
DB 251 GACATTATGTTATAGTTGTATGGAAC 278

RESULT 19
AAX78792
ID AAX78792 standard; DNA; 822 BP.
XX AC AAX78792;
XX XX 06-SEP-1999 (first entry)
XX DT

```

```

DE HPV fusion protein DI/3-E6-His/HPV16 DNA.
XX
XX Fusion protein; E6 protein; E7 protein; E6/E7; immunomodulator; tumour;
KM immunological fusion partner; CpG oligonucleotide; immune response;
KM HPV antigen; prevention; treatment; ss.
XX
OS Synthetic.
OS Human papillomavirus.
XX
XX WO9933868-A2.
XX
XX 08-JUL-1999.
XX
XX 18-DEC-1998; 98WO-EP008563.
XX
XX 24-DEC-1997; 97GB-00027262.
XX
XX (SMK ) SMITHKLINE BEECHAM BIOLOGICALS.
XX
XX Dalemans WJ, Gerard CMG;
XX
XX WPI, 1999-405485/34.
XX
XX P-PSDB; AAY5376.
XX
XX Composition comprising an E6, E7 or E6/E7 fusion protein from HPV to
PT induce immune response to HPV.
XX
XX Example II; Page 48; 62pp; English.
XX
XX AAY78791-X78801 represent nucleic acid sequences which encode novel
CC constructs comprising an E6 or E7 protein or E6/E7 fusion protein from
CC HPV (represented in AAY5375-Y25386). These constructs are optionally
CC linked to an immunological fusion partner and an immunomodulatory CpG
CC oligonucleotide. The products of the invention can be used to induce an
CC immune response in a patient to an HPV antigen. They can also be used for
CC preventing or treating HPV induced tumours
XX
XX Sequence 822 BP; 277 A; 147 C; 168 G; 230 T; 0 U; 0 Other;
SQ
Query Match 100.0%; Score 28; DB 2; Length 822;
Best Local Similarity 100.0%; Pred. No. 0.41;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 GACATTATTGTTATGTTGATGGAAC 28
DB 569 GACATTATTGTTATGTTGATGGAAC 596

```

```

XX
XX (SMK ) SMITHKLINE BEECHAM BIOLOGICALS.
XX
XX Bruck C, Cabezon Silva T, Delisse AEF, Gerard CMG;
XX Lombardo-Bencheikh A;
XX
XX WPI, 1999-190587/16.
XX
XX Human Papilloma Virus (HPV) fusion proteins - useful in vaccines for
PT treatment or prophylaxis of HPV induced lesions.
XX
XX Disclosure; Fig 3; 95pp; English.
XX
XX This sequence represents the coding region for a chimeric E6 or E7
CC protein or E6/E7 fusion protein from Human papillomavirus (HPV) linked to
CC an immunological fusion partner. In this case, a fragment of the
CC Haemophilus influenza B protein D. The sequence also contains a
CC histidine tag at the C-terminus of the encoded protein. The protein can
CC be used in a vaccine, for immuno-therapeutically treating HPV induced
CC tumour lesions (benign or malignant) and preventing HPV viral infection.
CC (Updated on 17-OCT-2003 to standardise OS field)
XX
XX Sequence 822 BP; 277 A; 147 C; 168 G; 230 T; 0 U; 0 Other;
SQ
Query Match 100.0%; Score 28; DB 2; Length 822;
Best Local Similarity 100.0%; Pred. No. 0.41;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 GACATTATTGTTATGTTGATGGAAC 28
DB 569 GACATTATTGTTATGTTGATGGAAC 596

```

```

RESULT 20
AAX29781
ID AAX29781 standard; DNA; 822 BP.
XX
XX AAX29781;
AC
XX 17-OCT-2003 (revised)
DT 22-JUN-1999 (first entry)
XX
DE Prot.DI/3-E6-His/HPV16 coding sequence.
XX
XX Chimeric; E6; E7; fusion protein; protein D; vaccine; immunotherapy;
KM tumour; lesion; benign; malignant; virus; infection; ss.
XX
OS Human papillomavirus.
OS Haemophilus influenzae.
OS Chimeric.
XX
XX WO9910375-A2.
XX
XX 04-MAR-1999.
XX
XX 17-AUG-1998; 98WO-EP005285.
XX
XX 22-AUG-1997; 97GB-00017953.

```

```

RESULT 21
AED52633
ID AED52633 standard; DNA; 822 BP.
XX
XX AED52633;
AC
XX 29-DEC-2005 (first entry)
DT
XX
DE Fusion protein DI/3-E6-His/HPV16, DNA.
XX
XX Fusion protein; vaccine; papilloma; cytostatic; papillomavirus infection;
KM viricide; uterine cervix tumor; E6; ds; gene; D protein.
XX
XX Haemophilus influenzae; strain 772.
OS Human papillomavirus type 16.
OS Synthetic.
OS Chimeric.
XX
XX Key Location/Qualifiers
FH 1..822
FT CDS /*tag=a
FT /product="Fusion protein DI/3-E6-His/HPV16"
XX
XX IN9801903-14.
XX
XX 04-MAR-2005.
XX
XX 24-AUG-1998; 98IN-CH001903.
XX
XX 22-AUG-1997; 97EP-00179535.
XX
XX (SMK ) SMITHKLINE BEECHAM BIOLOGICALS.
XX
XX Tyrrell AWR;
XX
XX WPI, 2005-557648/57.
XX
XX P-PSDB; AED52634.
XX
XX Vaccine.
XX

```

```

PS Example 4; Fig 3; 96pp; English.
CC The invention relates to human Papilloma virus (HPV) fusion proteins,
CC linked to an immunological fusion partner that provides a helper epitopes
CC to the HPV antigen. Vaccine formulations comprising the fusion proteins
CC are useful in the treatment or prophylaxis of HPV induced lesions
CC (papillomas, warts and cervical cancer). HPV 16 (and HPV 18) E6 and E7
CC proteins (singly, as an E6-E7 fusion or mutated) were fused to either
CC Haemophilus influenzae D protein (20-127), the C-terminus of
CC Streptococcus pneumoniae LysA protein (CLYA) or thiodoxin. The present
CC sequence encodes an HPV-H. influenzae D protein, fusion protein of the
CC invention.
SQ Sequence 822 BP; 277 A; 147 C; 168 G; 230 T; 0 U; 0 Other;

Query Match          100.0%; Score 28; DB 14; Length 822;
Best Local Similarity 100.0%; Pred. No. 0.41;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 GACATTATTGTTATAGTTGATGGAAC 28
Db 569 GACATTATTGTTATAGTTGATGGAAC 596

RESULT 22
AAK78795
ID AAK78795 standard; DNA; 879 BP.
AC AAK78795;
XX
XX
XX 06-SEP-1999 (first entry)
DT
XX
XX HPV fusion protein CLYTA-B6-His/HPV16 DNA.
DE
XX
XX Fusion protein; E6 protein; E7 protein; E6/E7; immunomodulator; tumour;
KW immunological fusion partner; Cpg oligonucleotide; immune response;
KW HPV antigen; prevention; treatment; ss.
XX
XX Synthetic.
OS Human papillomavirus.
XX
XX MO9933868-A2.
XX
XX 08-JUL-1999.
PD
XX
XX 18-DEC-1998; 98WO-EP008563.
PF
XX 24-DEC-1997; 97GB-00027262.
PR
XX (SMK ) SMITHKLINE BEECHAM BIOLOGICALS.
PA
XX Dalemans WLJ, Gerard CMG;
PI
XX MPI; 1999-405485/34.
DR P-PSDB; AAY25379.
XX
XX Composition comprising an E6, E7 or E6/E7 fusion protein from HPV to
PT induce immune response to HPV.
XX
XX Example VI; Page 52; 62pp; English.
XX
XX AAK78791-K78801 represent nucleic acid sequences which encode novel
CC constructs comprising an E6 or E7 protein or E6/E7 fusion protein from
CC HPV (represented in AAY5375-Y25386). These constructs are optionally
CC linked to an immunological fusion partner and an immunomodulatory Cpg
CC oligonucleotide. The products of the invention can be used to induce an
CC immune response in a patient to an HPV antigen. They can also be used for
CC preventing or treating HPV induced tumours
XX
XX Sequence 879 BP; 289 A; 168 C; 209 G; 213 T; 0 U; 0 Other;
SQ

Query Match          100.0%; Score 28; DB 2; Length 879;
Best Local Similarity 100.0%; Pred. No. 0.42;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 GACATTATTGTTATAGTTGATGGAAC 28
Db 626 GACATTATTGTTATAGTTGATGGAAC 653

RESULT 23
AAK29784
ID AAK29784 standard; DNA; 879 BP.
AC AAK29784;
XX
XX 17-OCT-2003 (revised)
DT 22-JUN-1999 (first entry)
XX
XX CLYTA-B6-His coding sequence.
DE
XX
XX Chimeric; E6; E7; fusion protein; CLYTA; vaccine; immunotherapy; tumour;
KW lesion; benign; malignant; virus; infection; ss.
XX
XX Human papillomavirus.
OS Streptococcus pneumoniae.
XX
XX Chimeric.
OS
XX MO9910375-A2.
XX
XX 04-MAR-1999.
PD
XX
XX 17-AUG-1998; 98WO-EP005285.
PF
XX 22-AUG-1997; 97GB-00017953.
PR
XX
XX (SMK ) SMITHKLINE BEECHAM BIOLOGICALS.
PA
XX Bruck C, Cabezon Silva T, Delisse AEF, Gerard CMG;
PI Lombardo-Bencheikh A;
XX
XX MPI; 1999-190587/16.
DR P-PSDB; AAY02635.
XX
XX Human Papilloma Virus (HPV) fusion proteins - useful in vaccines for
PT treatment or prophylaxis of HPV induced lesions.
XX
XX Disclosure; Fig 10; 95pp; English.
XX
XX This sequence represents the coding region for a chimeric E6 or E7
CC protein or E6/E7 fusion protein from Human papillomavirus (HPV) linked to
CC an immunological fusion partner, in this case, a fragment of the
CC Streptococcus pneumoniae CLYTA protein. The sequence also contains a
CC histidine tag at the C-terminus of the encoded protein. The protein can
CC be used in a vaccine, for immuno-therapeutically treating HPV induced
CC tumour lesions (benign or malignant) and preventing HPV viral infection.
CC (Updated on 17-OCT-2003 to standardise OS field)
XX
XX Sequence 879 BP; 289 A; 168 C; 209 G; 213 T; 0 U; 0 Other;
SQ

Query Match          100.0%; Score 28; DB 2; Length 879;
Best Local Similarity 100.0%; Pred. No. 0.42;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 GACATTATTGTTATAGTTGATGGAAC 28
Db 626 GACATTATTGTTATAGTTGATGGAAC 653

RESULT 24
AED52641
ID AED52641 standard; DNA; 879 BP.
AC AED52641;
XX
XX 29-DEC-2005 (first entry)
DT

```



```

XX DE Fusion protein clyta-E6-His/HPV16, DNA.
XX KM Fusion protein; vaccine; papilloma; cytostatic; papillomavirus infection;
XX KM virucide; uterine cervix tumor; E6; db; gene; lyta.
XX OS Streptococcus pneumoniae.
XX OS Human papillomavirus type 16.
XX OS Synthetic.
XX OS Chimeric.
XX FH Key
XX FT CDS
XX FT /tag= a
XX FT /product= "Fusion protein clyta-E6-His/HPV16"
XX IN9801903-14.
XX PD 04-MAR-2005.
XX PP 24-AUG-1998; 98IN-CH001903.
XX PR 22-AUG-1997; 97EP-00179535.
XX PA (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.
XX PI Tyrtrell AMR;
XX DR WPI; 2005-557648/57.
XX DR P-PSDB; AED52642.
XX PT Vaccine.
XX PS Example 10; Fig 10; 96pp; English.
XX CC The invention relates to human Papilloma virus (HPV) fusion proteins,
XX CC linked to an immunological fusion partner that provides T helper epitopes
XX CC to the HPV antigen. Vaccine formulations comprising the fusion proteins
XX CC are useful in the treatment or prophylaxis of HPV induced lesions
XX CC (papillomas, warts and cervical cancer). HPV 16 (and HPV 18) E6 and E7
XX CC proteins (bingly, as an E6-E7 fusion or mutated) were fused to either
XX CC Haemophilus influenzae D protein (20-127), the C-terminus of
XX CC Streptococcus pneumoniae Lyta protein (clyta) or thioresoxin. The present
XX CC sequence encodes an HPV-Lyta, fusion protein of the invention.
XX SQ Sequence 879 BP; 289 A; 168 C; 209 G; 213 T; 0 U; 0 Other;

Query Match 100.0%; Score 28; DB 14; Length 879;
Best Local Similarity 100.0%; Pred. No. 0.42;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACATTATTGTTATAGTTGTATGGAAC 28
DB 626 GACATTATTGTTATAGTTGTATGGAAC 653

RESULT 25
AAD35101
ID AAD35101 standard; DNA; 939 BP.
XX AC AAD35101;
XX DT 25-JUL-2002 (first entry)
XX DE Human papillomavirus enhB6.7 construct DNA.
XX KM Alphavirus vector system; human papilloma virus; HPV; cervical cancer;
XX KM therapy; vaccine; virucide; cytostatic; de.
XX OS Human papillomavirus.
XX EP1195438-A1.
XX PN
XX PA

```

```

PD 10-APR-2002.
XX PF 06-OCT-2000; 2000EP-00203472.
XX PR 06-OCT-2000; 2000EP-00203472.
XX PA (UYGR-) RIJKSUNIV GRONINGEN.
XX PI Regts DG, Wilschut JC, Holtrop M, Daemen CAHH;
XX DR WPI; 2002-354156/39.
XX PT New alphavirus system, useful for genetic immunization against cervical
XX PT cancer, comprises papilloma virus nucleic acid.
XX PS Example 2; Fig 19; 45pp; English.
XX CC The present invention relates to an alphavirus vector system comprising
XX CC nucleic acid derived from a human papilloma virus (HPV). The invention or
XX CC cells containing it, are used in treatment and prevention of cervical
XX CC cancer, particularly as a vaccine. By selecting the nucleic acid that
XX CC encode E6/E7 proteins without ability to bind to pRb and p53, the risk
XX CC that cells infected with the alphavirus vector system may become
XX CC oncogenic is avoided (contrast use of other viral vectors). The present
XX CC sequence is Human papillomavirus enhB6.7 construct DNA
XX SQ Sequence 939 BP; 284 A; 197 C; 217 G; 241 T; 0 U; 0 Other;

Query Match 100.0%; Score 28; DB 6; Length 939;
Best Local Similarity 100.0%; Pred. No. 0.42;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACATTATTGTTATAGTTGTATGGAAC 28
DB 415 GACATTATTGTTATAGTTGTATGGAAC 442

RESULT 26
AAQ08627
ID AAQ08627 standard; DNA; 1000 BP.
XX AC AAQ08627;
XX DT 24-OCT-2003 (revised)
XX DT 21-APR-1994 (first entry)
XX DE HPV-16 fragment.
XX KM HPV-18; HPV-16; amplification; primer; polymerase chain reaction; PCR;
XX KM ss.
XX OS Human papillomavirus type 16.
XX FH Key
XX FH misc_binding
XX FT 198.207
XX FT /tag= a
XX FT /note= "primer (AAQ08628) binding site"
XX FT /tag= b
XX FT /note= "primer (AAQ08629) binding site"
XX FT 658.677
XX FT /tag= c
XX FT /note= "primer (AAQ08630) binding site"
XX PN DE3838269-A.
XX PD 17-MAY-1990.
XX PF 11-NOV-1988; 88DE-03838269.
XX PR 11-NOV-1988; 88DE-03838269.
XX PA (BEHM ) BEHRINGWERKE AG.

```

```

XX Cerutti P, Whitcomb J, Zijlstra J, Devilliers EM;
XX WPI; 1990-156905/21.
XX
XX Detection of human papilloma virus - by DNA amplification and analysis.
XX
XX Example 4b; Page 4; 11pp; German.
XX
XX Example 4b describes the results of the amplification of HPV-16 DNA by
XX PCR using primers. (Updated on 24-OCT-2003 to standardise OS field)
XX
XX Sequence 1000 BP; 340 A; 171 C; 230 G; 259 T; 0 U; 0 Other;
XX
XX Query Match 100.0%; Score 28; DB 2; Length 1000;
XX Best Local Similarity 100.0%; Pred. No. 0.42;
XX Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 GACATTATTGTTATAGTTGTATGGAAC 28
DB 333 GACATTATTGTTATAGTTGTATGGAAC 360

RESULT 27
AAN91784
ID AAN91784 standard; DNA; 1005 BP.
XX
XX AAN91784;
XX
XX 27-AUG-2003 (revised)
XX 25-MAR-2003 (revised)
XX 16-MAR-1990 (first entry)
XX
XX DNA probe complementary to human papilloma virus type 16.
XX
XX Cervical cancer.
XX
XX
XX Human papillomavirus type 16.
XX
XX W08909940-A.
XX
XX 19-OCT-1989.
XX
XX 04-APR-1989; 89WO-US0013318.
XX
XX 04-APR-1988; 88US-00177404.
XX
XX 31-MAR-1989; 89US-00330381.
XX
XX (ONCO-) ONCOR INC.
XX
XX George AL, Groff DE;
XX
XX WPI; 1989-324314/44.
XX
XX Rapid detection of specific human papilloma virus genotypes - by
XX hybridisation of DNA digest with new labelled nucleic acid probes.
XX
XX Claim 40; Page 46; 81pp; English.
XX
XX Obtd. by cutting HPV16 with BamHI and PvuII. The patent describes probes
XX (DNA or RNA) and their complements capable of detecting one or a
XX combination of HPV types 6, 11, 16, 18, 31, 33 and 35. (Updated on 25-MAR
XX -2003 to correct PR field.) (Updated on 27-AUG-2003 to correct OS field.)
XX
XX Sequence 1005 BP; 325 A; 182 C; 190 G; 308 T; 0 U; 0 Other;
XX
XX Query Match 100.0%; Score 28; DB 1; Length 1005;
XX Best Local Similarity 100.0%; Pred. No. 0.42;
XX Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 GACATTATTGTTATAGTTGTATGGAAC 28
DB 785 GACATTATTGTTATAGTTGTATGGAAC 812

```

```

RESULT 28
AAK78793
ID AAK78793 standard; DNA; 1116 BP.
XX
XX AAK78793;
XX
XX 06-SEP-1999 (first entry)
XX
XX HPV fusion protein DI/3-E6E7-His/HPV16 DNA.
XX
XX Fusion protein; E6 protein; E7 protein; E6/E7; immunomodulator; tumour;
XX immunological fusion partner; CpG oligonucleotide; immune response;
XX HPV antigen; prevention; treatment; ss.
XX
XX Synthetic.
XX
XX Human papillomavirus.
XX
XX W0933868-A2.
XX
XX 08-JUL-1999.
XX
XX 18-DEC-1998; 98WO-EP008563.
XX
XX 24-DEC-1997; 97GB-00027262.
XX
XX (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.
XX
XX Dalemans WLJ, Gerard CMG;
XX
XX WPI; 1999-405485/34.
XX
XX P-PSDB; AAY25377.
XX
XX Composition comprising an E6, E7 or E6/E7 fusion protein from HPV to
XX induce immune response to HPV.
XX
XX
XX Example III; Page 49; 62pp; English.
XX
XX AAK78791-K78801 represent nucleic acid sequences which encode novel
XX constructs comprising an E6 or E7 protein or E6/E7 fusion protein from
XX HPV (represented in AAY25375-125386). These constructs are optionally
XX linked to an immunological fusion partner and an immunomodulatory CpG
XX oligonucleotide. The products of the invention can be used to induce an
XX immune response in a patient to an HPV antigen. They can also be used for
XX preventing or treating HPV induced tumours
XX
XX Sequence 1116 BP; 368 A; 208 C; 235 G; 305 T; 0 U; 0 Other;
XX
XX Query Match 100.0%; Score 28; DB 2; Length 1116;
XX Best Local Similarity 100.0%; Pred. No. 0.42;
XX Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 GACATTATTGTTATAGTTGTATGGAAC 28
DB 569 GACATTATTGTTATAGTTGTATGGAAC 596

RESULT 29
AAK29782
ID AAK29782 standard; DNA; 1116 BP.
XX
XX AAK29782;
XX
XX 17-OCT-2003 (revised)
XX 22-JUN-1999 (first entry)
XX
XX Proc.DI/3-E6-E7-His/HPV16 coding sequence.
XX
XX Chimeric; E6; E7; fusion protein; protein D; vaccine; immunotherapy;
XX tumour; lesion; benign; malignant; virus; infection; ss.
XX
XX Human papillomavirus.

```

```

OS Haemophilus influenzae.
OS Chimeric.
XX
XX WO9910375-A2.
XX
XX 04-MAR-1999.
XX
XX 17-AUG-1998; 98WO-EP005285.
XX
XX 22-AUG-1997; 97GB-00017953.
XX
XX (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.
XX
XX Bruck C, Cabezon Silva T, Delisse AEF, Gerard CMG;
XX Lombardo-Bencheikh A;
XX
XX WPI; 1999-190587/16.
XX
XX P-PSDB; AAY02633.
XX
XX Human Papilloma Virus (HPV) fusion proteins - useful in vaccines for
XX treatment or prophylaxis of HPV induced lesions.
XX
XX Disclosure; Fig 6; 95pp; English.
XX
XX This sequence represents the coding region for a chimeric E6 or E7
XX protein or E6/E7 fusion protein from Human papillomavirus (HPV) linked to
XX an immunological fusion partner, in this case, a fragment of the
XX Haemophilus influenzae B protein D. The sequence also contains a
XX histidine tag at the C-terminus of the encoded protein. The protein can
XX be used in a vaccine, for immuno-therapeutically treating HPV induced
XX tumour lesions (benign or malignant) and preventing HPV viral infection.
XX (Updated on 17-OCT-2003 to standardise OS field)
XX
XX Sequence 1116 BP; 368 A; 208 C; 235 G; 305 T; 0 U; 0 Other;
SQ
Query Match 100.0%; Score 28; DB 2; Length 1116;
Best Local Similarity 100.0%; Pred. No. 0.42;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACATTATTGTTATGTTGATGGAAC 28
Db 569 GACATTATTGTTATGTTGATGGAAC 596

RESULT 30
AED52637
ID AED52637 standard; DNA; 1116 BP.
XX
XX AED52637;
XX
XX 29-DEC-2005 (first entry)
XX
XX Fusion protein D1/3-E6-E7-His/HPV16, DNA.
XX
XX Fusion protein; vaccine; papilloma; cytostatic; papillomavirus infection;
XX virucide; uterine cervix tumor; E7; E6; ds; gene; D protein.
XX
XX Haemophilus influenzae; strain 772.
XX
XX Human papillomavirus type 16.
XX
XX Synthetic.
XX
XX Chimeric.
XX
XX Key Location/Qualifiers
XX CDS 1..1116
XX FT /*tag= a
XX FT /product= "Fusion protein D1/3-E6-E7-His/HPV16"
XX
XX IN9801903-14.
XX
XX 04-MAR-2005.
XX
XX 24-AUG-1998; 98IN-CH001903.
XX

```

```

PR 22-AUG-1997; 97EP-00179535.
XX
XX (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.
XX
XX Tyrrell AMR;
XX
XX WPI; 2005-557648/57.
XX
XX P-PSDB; AED52638.
XX
XX Vaccine.
XX
XX Example 6; Fig 6; 96pp; English.
XX
XX The invention relates to human Papilloma virus (HPV) fusion proteins,
XX linked to an immunological fusion partner that provides T helper epitopes
XX to the HPV antigen. Vaccine formulations comprising the fusion proteins
XX are useful in the treatment or prophylaxis of HPV induced lesions
XX (papillomas, warts and cervical cancer). HPV 16 (and HPV 18) E6 and E7
XX proteins (singly, as an E6-E7 fusion or mutated) were fused to either
XX Haemophilus influenzae D protein (20-127), the C-terminus of
XX Streptococcus pneumoniae Lyta protein (ClyA) or chlorodoxin. The present
XX sequence encodes an HPV-H. influenzae D protein, fusion protein of the
XX invention.
XX
XX Sequence 1116 BP; 368 A; 208 C; 235 G; 305 T; 0 U; 0 Other;
SQ
Query Match 100.0%; Score 28; DB 14; Length 1116;
Best Local Similarity 100.0%; Pred. No. 0.42;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACATTATTGTTATGTTGATGGAAC 28
Db 569 GACATTATTGTTATGTTGATGGAAC 596

RESULT 31
AAK78797
ID AAK78797 standard; DNA; 1173 BP.
XX
XX AAK78797;
XX
XX 06-SEP-1999 (first entry)
XX
XX HPV fusion protein CLYTA-E6E7-His/HPV16 DNA.
XX
XX Fusion protein; E6 protein; E7 protein; E6/E7; immunomodulator; tumour;
XX immunological fusion partner; CpG oligonucleotide; immune response;
XX HPV antigen; prevention; treatment; ss.
XX
XX Synthetic.
XX
XX Human papillomavirus.
XX
XX WO9933868-A2.
XX
XX 08-JUL-1999.
XX
XX 18-DEC-1998; 98WO-EP008563.
XX
XX 24-DEC-1997; 97GB-00027262.
XX
XX (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.
XX
XX Dalemans WLJ, Gerard CMG;
XX
XX WPI; 1999-405485/34.
XX
XX P-PSDB; AAY25381.
XX
XX Composition comprising an E6, E7 or E6/E7 fusion protein from HPV to
XX induce immune response to HPV.
XX
XX Example VIII; Page 54-55; 62pp; English.
XX
XX AAK78791-X78801 represent nucleic acid sequences which encode novel
XX

```

CC constructs comprising an E6 or E7 protein or E6/E7 fusion protein from
 CC HPV (represented in AAY25375-Y25386). These constructs are optionally
 CC linked to an immunological fusion partner and an immunomodulatory CpG
 CC oligonucleotide. The products of the invention can be used to induce an
 CC immune response in a patient to an HPV antigen. They can also be used for
 CC preventing or treating HPV induced tumours

XX Sequence 1173 BP; 380 A; 229 C; 276 G; 288 T; 0 U; 0 Other;
 SQ

Query Match 100.0%; Score 28; DB 2; Length 1173;
 Best Local Similarity 100.0%; Pred. No. 0.42;
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACATTATTGTTATGTTGTATGGAAC 28
 Db 626 GACATTATTGTTATGTTGTATGGAAC 653

RESULT 32
 AAY29786
 ID AAY29786 standard; DNA; 1173 BP.
 XX
 AC AAY29786;
 XX
 DT 17-OCT-2003 (revised)
 DT 22-JUN-1999 (first entry)
 XX
 XX CLYTA-E6E7-His coding sequence.
 DE
 XX Chimeric; E6; E7; fusion protein; CLYTA; vaccine; immunotherapy; tumour;
 KM lesion; benign; malignant; virus; infection; ss.
 XX
 OS Human papillomavirus.
 OS Streptococcus pneumoniae.
 OS Chimeric.
 XX
 PN MO9910375-A2.
 XX
 PD 04-MAR-1999.
 PD
 PF 17-AUG-1998; 98WO-EP005285.
 PF
 PR 22-AUG-1997; 97GB-00017953.
 PR
 PA (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
 PA
 PI Bruck C, Cabezon Silva T, Delisse AEF, Gerard CMG;
 PI Lombardo-Bencheikh A;
 PI
 XX WPI; 1999-190587/16.
 DR P-PSDB; AAY02637.
 DR
 PT Human Papilloma Virus (HPV) fusion proteins - useful in vaccines for
 PT treatment or prophylaxis of HPV induced lesions.
 PT
 XX
 PS Disclosure; Fig 14; 95pp; English.
 PS
 XX This sequence represents the coding region for a chimeric E6 or E7
 CC protein or E6/E7 fusion protein from Human papillomavirus (HPV) linked to
 CC an immunological fusion partner, in this case, a fragment of the
 CC Streptococcus pneumoniae CLYTA protein. The sequence also contains a
 CC histidine tag at the C-terminus of the encoded protein. The protein can
 CC be used in a vaccine, for immuno-therapeutically treating HPV induced
 CC tumour lesions (benign or malignant) and preventing HPV viral infection.
 CC (Updated on 17-OCT-2003 to standardise OS field)
 CC
 XX
 SQ Sequence 1173 BP; 380 A; 229 C; 276 G; 288 T; 0 U; 0 Other;
 SQ

Query Match 100.0%; Score 28; DB 2; Length 1173;
 Best Local Similarity 100.0%; Pred. No. 0.42;
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACATTATTGTTATGTTGTATGGAAC 28
 Db 626 GACATTATTGTTATGTTGTATGGAAC 28

Db 626 GACATTATTGTTATGTTGTATGGAAC 653

RESULT 33
 AED52645
 ID AED52645 standard; DNA; 1173 BP.
 XX
 AC AED52645;
 XX
 DT 29-DEC-2005 (first entry)
 DT
 DE Fusion protein cLYTA-E6-E7-His/HPV16, DNA.
 DE
 XX Fusion protein; vaccine; papilloma; cytostatic; papillomavirus infection;
 KM virucide; uterine cervix tumor; E7; E6; de; gene; Lyta.
 XX
 XX Streptococcus pneumoniae.
 OS
 OS Human papillomavirus type 16.
 OS
 OS Synthetic.
 OS Chimeric.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..1173
 FT /*tag= a
 FT /product= "Fusion protein cLYTA-E6-E7-His/HPV16"
 XX
 FN IN9801903-14.
 FN
 PD 04-MAR-2005.
 PD
 PF 24-AUG-1998; 98IN-CH001903.
 PF
 PR 22-AUG-1997; 97EP-00179535.
 PR
 PA (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
 PA
 PI Tyrrell AMR;
 PI
 XX WPI; 2005-557648/57.
 DR P-PSDB; AED52646.
 DR
 PT Vaccine.
 PT
 XX
 PS Example 12; Fig 14; 96pp; English.
 PS
 XX The invention relates to human Papilloma virus (HPV) fusion proteins,
 CC linked to an immunological fusion partner that provides T helper epitopes
 CC to the HPV antigen. Vaccine formulations comprising the fusion proteins
 CC are useful in the treatment or prophylaxis of HPV induced lesions
 CC (papillomas, warts and cervical cancer). HPV 16 (and HPV 18) E6 and E7
 CC proteins (singly, as an E6-E7 fusion or mutated) were fused to either
 CC Haemophilus influenzae D protein (20-127), the C-terminus of
 CC Streptococcus pneumoniae Lyta protein (cLYTA) or thioredoxin. The present
 CC sequence encodes an HPV-LyTA, fusion protein of the invention.
 CC
 XX
 SQ Sequence 1173 BP; 380 A; 229 C; 276 G; 288 T; 0 U; 0 Other;
 SQ

Query Match 100.0%; Score 28; DB 14; Length 1173;
 Best Local Similarity 100.0%; Pred. No. 0.42;
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACATTATTGTTATGTTGTATGGAAC 28
 Db 626 GACATTATTGTTATGTTGTATGGAAC 653

RESULT 34
 AAF55127
 ID AAF55127 standard; DNA; 7840 BP.
 AC AAF55127;
 XX

DT 29-MAY-2001 (first entry)
XX Nucleotide sequence of the vector pRetROFF-BEG7.
XX Stem cell; gene therapy; cell therapy; stem cell disorder; ss.
XX Synthetic.
XX
FH Key
FT misc_feature
FT Location/Qualifiers
FT 5600
FT /*tag= d
FT /note= "this base is given as N because it was illegible
FT in the figure"
FT 5650
FT /*tag= e
FT /note= "this base is given as N because it was illegible
FT in the figure"
FT 5700
FT /*tag= f
FT /note= "this base is given as N because it was illegible
FT in the figure"
FT 5750
FT /*tag= g
FT /note= "this base is given as N because it was illegible
FT in the figure"
FT 5800
FT /*tag= h
FT /note= "this base is given as N because it was illegible
FT in the figure"
FT 5850
FT /*tag= i
FT /note= "this base is given as N because it was illegible
FT in the figure"
FT 5850
FT /*tag= a
FT /note= "this base is given as N because it was illegible
FT in the figure"
FT 5900
FT /*tag= j
FT /note= "this base is given as N because it was illegible
FT in the figure"
FT 5900
FT /*tag= b
FT /note= "this base is given as N because it was illegible
FT in the figure"
FT 5950
FT /*tag= k
FT /note= "this base is given as N because it was illegible
FT in the figure"
FT 5950
FT /*tag= c
FT /note= "this base is given as N because it was illegible
FT in the figure"
FT 6000
FT /*tag= l
FT /note= "this base is given as N because it was illegible
FT in the figure"
FT 6050
FT /*tag= m
FT /note= "this base is given as N because it was illegible
FT in the figure"
FT 6100
FT /*tag= n
FT /note= "this base is given as N because it was illegible
FT in the figure"
FT 6150
FT /*tag= o
FT /note= "this base is given as N because it was illegible
FT in the figure"
FT 6200
FT /*tag= p
FT /note= "this base is given as N because it was illegible
FT in the figure"

FT misc_feature
FT 6250
FT /*tag= q
FT /note= "this base is given as N because it was illegible
FT in the figure"
FT 6300
FT /*tag= r
FT /note= "this base is given as N because it was illegible
FT in the figure"
FT 6350
FT /*tag= s
FT /note= "this base is given as N because it was illegible
FT in the figure"
FT 6400
FT /*tag= t
FT /note= "this base is given as N because it was illegible
FT in the figure"
FT 6450
FT /*tag= u
FT /note= "this base is given as N because it was illegible
FT in the figure"
FT 6500
FT /*tag= v
FT /note= "this base is given as N because it was illegible
FT in the figure"
FT 6600
FT /*tag= w
FT /note= "this base is given as N because it was illegible
FT in the figure"
FT 6650
FT /*tag= x
FT /note= "this base is given as N because it was illegible
FT in the figure"
FT 6700
FT /*tag= y
FT /note= "this base is given as N because it was illegible
FT in the figure"
FT 6750
FT /*tag= z
FT /note= "this base is given as N because it was illegible
FT in the figure"
FT 6800
FT /*tag= aa
FT /note= "this base is given as N because it was illegible
FT in the figure"
FT 6850
FT /*tag= ab
FT /note= "this base is given as N because it was illegible
FT in the figure"
FT 6900
FT /*tag= ae
FT /note= "this base is given as N because it was illegible
FT in the figure"
FT 6950
FT /*tag= af
FT /note= "this base is given as N because it was illegible
FT in the figure"
FT 7000
FT /*tag= ag
FT /note= "this base is given as N because it was illegible
FT in the figure"
FT 7050
FT /*tag= ah
FT /note= "this base is given as N because it was illegible
FT in the figure"
FT 7100
FT /*tag= ai
FT /note= "this base is given as N because it was illegible
FT in the figure"
FT 7150
FT /*tag= aj
FT /note= "this base is given as N because it was illegible
FT in the figure"
FT 7200
FT misc_feature


```

FT CDS 562..858
FT /tag= e
FT /product= "E7 transforming protein"
FT /note= "E7 ORF 544 to 858; putative"
FT CDS 865..2813
FT /tag= f
FT /product= "E1 replication protein"
FT /note= "E1 interrupted ORF from 859 to 2813; putative"
FT CDS 2755..3852
FT /tag= g
FT /product= "E2 regulatory protein"
FT /note= "E2 ORF from 2725 to 3852; putative"
FT CDS 3332..3619
FT /tag= h
FT /product= "E4"
FT /note= "E4 ORF from 3332 to 3619; putative"
FT CDS 3863..4099
FT /tag= i
FT /product= "E5"
FT /note= "E5 ORF from 3863 to 4099; putative"
FT polyA_signal 4213..4218
FT /tag= j
FT /note= "Putative"
FT CDS 4235..5656
FT /tag= k
FT /product= "L2 minor capsid protein"
FT /note= "L2 ORF from 4133 to 5656; putative"
FT TATA_signal 4289..4295
FT /tag= l
FT /product= "L1 major capsid protein"
FT /note= "L1 ORF from 5526 to 7154; putative"
FT polyA_signal 7260..7265
FT /tag= n
FT US5679509-A.
FT 21-OCT-1997.
FT 30-SEP-1994; 94US-00316239.
FT 28-SEP-1993; 93US-00127906.
FT (VUNE-) UNIV NEW MEXICO STATE.
FT Wheeler CM, Parmenter CA;
FT WPI; 1997-525714/48.
FT P-PSDB; AAM35741.
FT Evaluating risk of cervical dysplasia or cervical cancer - by detecting
FT variant form of human papilloma virus 16.
FT Claim 1; Col 9-16; 33pp; English.
XX
CC Methods have been developed for distinguishing a subset of human
CC papilloma virus (HPV) that is associated with an increased risk of
CC developing cervical dysplasia or cervical cancer. The methods involve:
CC (1) preparing a cervical sample to expose any HPV-16 B6 gene in the
CC sample and determining if the base at position 350 of the B6 gene (see
CC AAT94723 and AAT94724 for comparison) is T or G, where the presence of G
CC at position 350 is associated with an increased risk of developing
CC cervical dysplasia or cervical cancer; and (2) preparing a cervical
CC sample to expose any HPV-16 B6 protein in the sample and determining if
CC the amino acid at position 83 of the protein (see position 90 in AAM35741
CC and AAM35742 for comparison) is Val or Leu, where the presence of Val at
CC position 83 that is associated with an increased risk of developing
CC cervical dysplasia or cervical cancer. The present sequence represents
CC the reference nucleotide sequence for HPV-16 B6. The 350G variant
CC correlates well with Pap scores; 350T:350G ratios among 45 HPV16 samples
CC were 10:4 for negative Pap scores; 4:2 for CIN I, 1:6 for CIN II; 2:9 for
CC CIN III; 0:3 for cancer. (Updated on 25-MAR-2003 to correct PF field.)

```

```

XX SQ Sequence 7904 BP; 2601 A; 1377 C; 1509 G; 2417 T; 0 U; 0 Other;
XX Query Match 100.0%; Score 28; DB 2; Length 7904;
XX Best Local Similarity 100.0%; Pred. No. 0.42;
XX Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 GACATTATTGTTATAGTTTGATGGAAC 28
DB 333 GACATTATTGTTATAGTTTGATGGAAC 360
XX
XX RESULT 37
XX ID AAX33881 standard; DNA; 7904 BP.
XX AAX33881;
XX AC 17-OCT-2003 (revised)
XX DT 25-JUN-1999 (first entry)
XX DE HPV-16 genomic sequence.
XX KW HPV-16; inhibitor; antisense oligonucleotide; E6/E7 gene; human;
XX KW keratinocyte; cervical cell; cervical tumour; PCR primer; beta-actin; ss.
XX OS Human papillomavirus type 16.
XX MO9913071-A1.
XX PD 18-MAR-1999.
XX PF 03-SEP-1998; 98WO-US018320.
XX PR 05-SEP-1997; 97US-00929140.
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX DiPaolo J, Alvarez-Salas L;
XX WPI; 1999-243727/20.
XX
XX New antisense oligonucleotide analogs for inhibiting growth of cervical
XX tumors.
XX Disclosure; Page 28-32; 40pp; English.
XX
XX This sequence represents a sequence from the HPV-16 genome. This
XX invention relates to antisense oligonucleotide analogs (ONs) that have a
XX sequence complementary to a sequence of nucleotides 415-445 of human
XX papilloma virus-16 (HPV-16). The antisense ONs can be used to inhibit
XX expression of HPV gene E6/E7 in living cells, preferably human
XX keratinocytes or human cervical cells. They bind to E6/E7 mRNA in the
XX cell, prevent mRNA translation and promote mRNA degradation by
XX intracellular RNase H. They can be used for preventing transformation of
XX living cells by HPV. The antisense ONs are used particularly for
XX inhibiting the growth of cervical tumours. (Updated on 17-OCT-2003 to
XX standardise OS field)
XX
XX SQ Sequence 7904 BP; 2601 A; 1377 C; 1509 G; 2417 T; 0 U; 0 Other;
XX Query Match 100.0%; Score 28; DB 2; Length 7904;
XX Best Local Similarity 100.0%; Pred. No. 0.42;
XX Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 GACATTATTGTTATAGTTTGATGGAAC 28
DB 333 GACATTATTGTTATAGTTTGATGGAAC 360
XX
XX RESULT 38
XX AAV99946
XX ID AAV99946 standard; DNA; 7904 BP.

```

XX	AAV999946;
AC	
XX	
DT	26-APR-1999 (first entry)
XX	
DE	Nucleotide sequence of the human papillomavirus 16.
XX	
KW	Human; involucrin; hINV; promoter; epithelial cell; oncogene;
XX	suprabasal epithelial cell; human papillomavirus; HPV; transgenic animal;
KW	hyperplasia; cancer; carcinogen; ds.
XX	
OS	Human papillomavirus.
XX	
PN	W09842826-A1.
XX	
PD	01-OCT-1998.
XX	
PF	24-MAR-1998; 98WO-US005866.
XX	
PR	24-MAR-1997; 97US-00822509.
XX	
PA	(RESE) RESEARCH CORP TECHNOLOGIES INC.
XX	
PI	Crieth JF, Eckert RL;
XX	
DR	WPI; 1999-070069/06.
XX	
PT	New isolated promoter sequences of human involucrin gene - used for
PT	specific expression of nucleic acids in epithelial cells, particularly
PT	for producing transgenic animals for use as models for cancers.
XX	
P8	Disclosure; Fig 2; 74pp; English.
XX	
CC	This is the nucleotide sequence of the genome of the Human papillomavirus
CC	16 used in the method of the invention involving human involucrin (hINV),
CC	in the production of a nucleotide sequence comprising the promoter
CC	sequences of the hINV gene. It can be used to direct the expression of
CC	nucleic acids in epithelial cells, particularly suprabasal epithelial
CC	cells. The nucleotide sequence can be used with human papillomavirus
CC	(HPV) oncogenes for producing transgenic animals which display cervical
CC	and epidermal hyperplasias as well as cancer of the trachea, esophagus,
CC	colon, epidermis, anus/rectum, lymph nodes, spleen and lung. These
CC	animals provide useful models for the identification of potential
CC	carcinogens and co-carcinogens, identification of anti-neoplastic
CC	compounds, identification of genes which play a role in neoplastic
CC	progression of cancers of the trachea, esophagus, colon, epidermis,
CC	anus/rectum, lymph nodes, spleen, lung and cervix. The products can also
CC	be used for producing antisense molecules which can be used for reducing
CC	or preventing expression of a gene whose expression is controlled by the
CC	2473 bp sequence
XX	
SO	Sequence 7904 BP; 2601 A; 1379 C; 1507 G; 2417 T; 0 U; 0 Other;
XX	
Query Match	100.0%; Score 28; DB 2; Length 7904;
Best Local Similarity	100.0%; Pred. No. 0.42;
Matches	28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Cy	1 GACATTATGTTAGTTCGATGGAAC 28
DB	333 GACATTATGTTAGTTCGATGGAAC 360
XX	
RESULT 39	
AAV99952	
ID	AAV99952 standard; DNA; 7904 BP.
AC	AAV99952;
XX	
DT	26-APR-1999 (first entry)
XX	
DE	Nucleotide sequence of the human papillomavirus 16.
XX	
KW	Human; involucrin; hINV; transgenic non-human; promoter; cancer;

KW epithelial cell; suprabasal epithelial cell; oncogene;
 KW epidermal hyperplasia; human papillomavirus; da.
 XX
 XX Human papillomavirus.
 OS
 XX
 XX MO9842825-A1.
 PN
 PD 01-OCT-1998.
 XX
 XX
 XX 24-MAR-1998; 98WO-US0005588.
 PE
 XX 24-MAR-1997; 97US-00822509.
 PR
 XX 24-MAR-1997; 97US-00822509.
 XX
 XX (UYCA-) UNIV CASE WESTERN RESERVE.
 PA (UYOH-) UNIV OHIO.
 XX
 XX
 PI Irish JF, Eckert RL, Wagner TE;
 XX
 XX WPI; 1999-070068/06.
 DR
 XX
 XX
 PT New transgenic non-human animals - contain nucleic acid sequences under
 PT the control of the human involucrin gene promoter, used for producing
 PT cancer models.
 XX
 XX
 PS Disclosure; Fig 2; 77pp; English.
 XX
 XX This is the nucleotide sequence of the human papillomavirus 16, involved
 CC in the method of the invention where the involucrin (hINV) gene is
 CC expressed by a transgenic non-human animal. The nucleotide sequence
 CC comprises the promoter of the hINV gene which directs the expression of
 CC nucleic acids in epithelial cells, particularly suprabasal epithelial
 CC cells. In particular when the NS directs the expression of oncogenes such
 CC as human papillomavirus (HPV) 16 oncogenes in transgenic animals, the
 CC animals display cervical and epidermal hyperplasias as well as cancer of
 CC the trachea, esophagus, colon, epidermis, anus/rectum, lymph nodes,
 CC spleen and lung. These animals provide useful models for the
 CC identification of potential carcinogens and co-carcinogens,
 CC identification of anti-neoplastic compounds, identification of genes
 CC which play a role in neoplastic progression of cancers of the trachea,
 CC esophagus, colon, epidermis, anus/rectum, lymph nodes, spleen, lung, and
 CC cervix
 XX
 XX
 SQ Sequence 7904 BP; 2601 A; 1379 C; 1507 G; 2417 T; 0 U; 0 Other;

 Query Match 100.0%; Score 28; DB 2; Length 7904;
 Best Local Similarity 100.0%; Pred. No. 0.42;
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0

 QY 1 GACATTATTGTTAGTTGATGGAAC 28
 ||||||||||||||||||||
 Db 333 GACATTATTGTTAGTTGATGGAAC 360

 RESULT 40
 ID ADA03127
 XX ADA03127 standard; DNA; 7904 BP.
 AC
 XX ADA03127;
 XX
 DT 06-NOV-2003 (first entry)
 XX
 DE HPV E6 sequence for inhibitor oligonucleotide repression method.
 XX
 KW cyostatic; vinuclide; anti-HIV; neuroprotective; ophthalmological;
 KW antidiabetic; antiporiatic; antirheumatic; antiathritic; da;
 KW inhibitor repression; inhibitor RNA; apoptosis; necrosis;
 KW differentiation; tumour cell division; BCL2 gene family;
 KW matrix metalloproteinase; membrane metalloproteinase;
 KW nuclear hormone receptor; transcription factor;
 KW vascular endothelial growth factor; p53; chromosomal translocation;
 KW leukaemia; androgen receptor; diagnosis; cancer; human papilloma virus;
 KW AIDS; BSE; CJD; macular degeneration; angiogenesis; diabetic retinopathy;
 KW psoriasis; rheumatoid arthritis.

XX OS Human papillomavirus.
 XX PN W02003040366-A2.
 XX PD 15-MAY-2003.
 XX PF 08-NOV-2002; 2002MO-FR003843.
 XX PR 09-NOV-2001; 2001FR-00014549.
 XX PR 10-APR-2002; 2002FR-00004474.
 XX PA (CNRS) CNRS CENT NAT RECH SCI.
 XX PI Harel-Bellan A, Alt-Si-Ali S, Cabon-Georget F, Chauchereau A,
 XX PI Dautry F;
 XX DR WPI; 2003-441571/41.
 XX PT New double-stranded oligonucleotides, useful e.g. for treatment and
 XX PT diagnosis of tumors, comprise complementary strands with single-stranded
 XX PT overhangs.
 XX PS Disclosure; Fig 6A; 148bp; French.
 XX CC The invention relates to double-stranded oligonucleotides consisting of
 CC two complementary strands (1a, 1b), each having 1-5 unpaired nucleotides,
 CC at either at their 3' and 5' ends, forming single-stranded overhangs. One
 CC of (1a) and (1b) is complementary to a target sequence, (DNA or RNA),
 CC that is to be specifically repressed. The oligonucleotides are preferably
 CC double stranded inhibitor RNA molecules with a couple of thymidine bases
 CC attached at the 3' or 5' ends. The targets are preferably nucleic acids
 CC that, when repressed, induce apoptosis, necrosis or differentiation of
 CC tumour cells and/or inhibit division of such cells. Typical of many
 CC specified targets include: genes of the BCL2 family; genes that encode
 CC metalloproteases (matrix or membrane) or their inhibitors; genes encoding
 CC mutant forms of nuclear hormone receptors; a sequence encoding the Hfl-
 CC alpha transcription factor; sequences encoding various isoforms of the
 CC vacuolar endothelial growth factor; viral genes; genes that express a
 CC mutated protein, e.g. inactive p53; genes that are formed by a
 CC chromosomal translocation, e.g. where associated with leukaemia, and
 CC genes that express androgen receptors. The oligonucleotides are used: (1)
 CC to study gene function; (11) for therapy or diagnosis, particularly of
 CC conditions caused by expression of a harmful gene or fusion protein,
 CC specifically cancer (e.g. associated with expression of mutant p53 or of
 CC the human papilloma virus B6 protein), viral infections, especially AIDS
 CC or cancer-inducing viruses, or unconventional infections, e.g. BSE or CJD
 CC ; and (111) for treating hypervascular diseases, e.g. age-related macular
 CC degeneration, angiogenesis in tumours, diabetic retinopathy, psoriasis
 CC and rheumatoid arthritis. They may also be used in vitro, e.g. for
 CC treating transplants and for establishing a genetic profile, for
 CC individualization, or modification, of treatment regimes. They provide
 CC very effective and very specific repression of genes. RNA hybrids are
 CC more stable than either hybrids prepared from DNA or single-stranded
 CC sequences; contain only natural components (so will not induce
 CC immunological or intolerance reactions) and they enter tumour cells more
 CC effectively than plasmids. This sequence represents the human
 CC papillomavirus B6 gene to which oligonucleotides can be targeted.
 XX CC
 SQ Sequence 7904 BP; 2601 A; 1377 C; 1509 G; 2417 T; 0 U; 0 Other;
 Query Match 100.0%; Score 28; DB 8; Length 7904;
 Best Local Similarity 100.0%; Pred. No. 0.42; Mismatches 0; Indels 0; Gaps 0;
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 GACATTATGTTATGTTGATGGAAC 28
 |||||
 ID ADZ45647 333 GACATTATGTTATGTTGATGGAAC 360
 |||||
 RESULT 41
 ADOS8335
 ID ADO58335 standard; DNA; 7904 BP.

XX AC ADO58335;
 XX DT 15-JUL-2004 (first entry)
 XX DE Human papillomavirus type 16 L1 capsid protein DNA.
 XX KW gene delivery system; virus-like particle; viral structure protein;
 KW capsid protein; L1; human papillomavirus; HPV; immunity; cytokine;
 KW chemokine; interleukin-2; IL-2; tumour suppressor gene H19; cell cycle;
 KW p21; angiogenesis inhibitor; flk-1; capsomer; HPV16; gene; ds.
 XX OS
 XX PN Human papillomavirus type 16.
 XX PN KR2002084620-A.
 XX PD 09-NOV-2002.
 XX PF 03-MAY-2001; 2001KR-00024124.
 XX PR 03-MAY-2001; 2001KR-00024124.
 XX PA (OHYK/) OH Y K.
 XX PI Kim JG, Ko JY, Oh YK, Shin GS, Son TY;
 XX PI WPI; 2003-338995/32.
 XX PT Gene delivery system containing desired foreign gene in virus-like
 XX PT particles.
 XX PS Example 1; SEQ ID NO 1; 16bp; Korean.
 XX CC The invention describes a gene delivery system containing a desired
 CC foreign gene in virus-like particles. The desired foreign gene can be
 CC stably delivered into a cell without being decomposed by nucleases and
 CC the desired foreign gene also stable in vivo because the virus-like
 CC particles do not have nucleic acids derived from virus. The gene delivery
 CC system contains a desired foreign gene in virus-like particles consisting
 CC of only viral structure proteins, wherein the viral structure protein is
 CC a capsid protein L1 derived from human papillomavirus (HPV); the foreign
 CC gene is associated with immunity; and the foreign gene associated with
 CC immunity is selected from the group consisting of cytokine gene,
 CC chemokine gene, interleukin-2 (IL-2) gene, tumour suppressor gene H19,
 CC cell cycle participating gene p21 and angiogenesis inhibitor flk-1. The
 CC method for producing the gene delivery system containing the desired
 CC foreign gene comprises the steps of: treating the virus-like particles
 CC with an appropriate reducing agent to be unit capsomer proteins; adding
 CC the desired foreign gene into the unit capsomer protein and allowing them
 CC to react; and removing the reducing agent to refold the unit capsomer
 CC protein. This sequence represents human papillomavirus type 16 (HPV16) L1
 CC capsid protein DNA.
 XX CC
 SQ Sequence 7904 BP; 2601 A; 1377 C; 1509 G; 2417 T; 0 U; 0 Other;
 Query Match 100.0%; Score 28; DB 11; Length 7904;
 Best Local Similarity 100.0%; Pred. No. 0.42; Mismatches 0; Indels 0; Gaps 0;
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 GACATTATGTTATGTTGATGGAAC 28
 |||||
 ID ADZ45647 333 GACATTATGTTATGTTGATGGAAC 360
 |||||
 RESULT 42
 ADZ45647
 ID ADZ45647 standard; DNA; 7904 BP.
 XX AC ADZ45647;
 XX DT 30-JUN-2005 (first entry)
 XX DE Human papillomavirus type 16 full length genomic clone - SEQ ID 1.

```

XX tumor marker; genome; diagnosis; cancer; uterine cervix tumor;
KM rectal tumor; colon tumor; ds.
XX
XX Human papillomavirus type 16.
XX WO200503333-A2.
XX
XX 14-APR-2005.
XX
XX 04-OCT-2004; 2004WO-DK000670.
XX
XX 07-OCT-2003; 2003DK-00001474.
XX 07-OCT-2003; 2003US-0509205P.
XX 13-FEB-2004; 2004US-0543925P.
XX
XX (DAKO-) DAKOCYTOMATION DENMARK AS.
XX
XX Chau MF, Bisgaard-Franzen K, Lin J, Rasmussen OF, Wang Z, Lusk J,
PI Lindberg M, Yost S;
XX
XX WPI; 2005-265441/29.
XX
XX New composition having a nucleic acid molecule identical to genomic clone
PT of human papilloma virus 16, 18 or 51, useful in diagnosing cancer or the
PT risk of developing cancer, in particular cervical, anal, colon and HPV-
PT related cancer.
XX
XX Claim 1; SEQ ID NO 1; 169pp; English.
XX
XX The invention comprises a composition for the detection of cancer
CC markers, the composition contains a nucleic acid molecule which is
CC substantially identical to a full length genomic clone of a human
CC papilloma virus (type 11, 16, 18, 51, 56, 58, 66, 70, or 73). The
CC composition of the invention is useful for diagnosing cancer or the risk
CC of developing cancer, in particular cervical, anal, colon and HPV-related
CC cancer. The present DNA sequence represents a human papilloma virus full
CC length genomic clone of the invention.
XX
XX Sequence 7904 BP; 2601 A; 1377 C; 1509 G; 2417 T; 0 U; 0 Other;
SQ
Query Match 100.0%; Score 28; DB 14; Length 7904;
Best Local Similarity 100.0%; Pred. No. 0.42;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GACATATTGTTATGTTGATGGAAC 28
DB 333 GACATATTGTTATGTTGATGGAAC 360

```

```

XX 08-JAN-2001; 2001US-00756095.
XX 08-JAN-2001; 2001US-00756096.
XX 08-JAN-2001; 2001US-00756097.
XX 20-APR-2001; 2001US-00838858.
XX 29-AUG-2001; 2001US-00941492.
XX
XX (INTR-) INTRON INC.
XX
XX Mitchell IG, Garcia-Bianco MA, Baker CC, Puttaraju M;
PI Mansfield GS, Chao H;
XX
XX WPI; 2002-566693/60.
XX
XX Novel cell having pre-trans-splicing molecules with target binding
PT domains that target binding of PTM to pre-mRNA, 3' or 5' splice region,
PT spacer region, nucleotide sequence to be trans-spliced to target-pre-
PT mRNA.
XX
XX Example; Fig 52; 229pp; English.
XX
XX The present invention describes a cell (I) comprising pre-trans-splicing
CC molecules (PTMs) (II) which have one or more target binding domains (IIs)
CC that target binding of PTM to pre-mRNA, 3' splice region (IId) that
CC includes branch point pyrimidine tract and 3' splice acceptor site, or 5'
CC splice site (IId), spacer region (IId) that separates RNA splice site
CC from target binding domain, and nucleotide sequence to (IId) be trans-
CC spliced to target-pre-mRNA. Optionally, the cell comprises (II) either
CC comprising: (A) (IId) and (IId) or (B) (IId) and (IId). The cell
CC may comprise a recombinant vector expressing (II). (I) has cytostatic,
CC immunosuppressive and antimicrobial activities, and can be used in gene
CC therapy. (II) comprising one or more (preferably two or more) (IId) and
CC (IId) (or (IId)), (IId) and (IId), or (II) comprising either (A) or (B)
CC (excluding (IId)), is useful for producing a chimeric RNA molecule in a
CC cell which involves contacting a target pre-mRNA expressed in the cell
CC with (II) that is recognised by nuclear splicing components. The chimeric
CC RNA produced comprises sequences encoding a toxin or translatable
CC protein. The nucleotide sequence to be trans-spliced to target pre-mRNA
CC preferably comprises nucleotide sequences comprising exons 1-10 of cyclic
CC fibrosis trans-membrane conductance regulator (CFTR). The chimeric RNA
CC molecule produced using (II) which either comprises (A) or (B) further
CC comprises a nucleotide sequence tag. (I) can be used for gene regulation,
CC gene repair and targeted cell death. (I) can be used for the treatment of
CC various diseases including genetic, infectious or autoimmune diseases and
CC proliferative disorders such as cancer and to regulate gene expression in
CC plants. ABQ73414 to ABQ73536 represent sequences used in the
CC exemplification of the present invention
XX
XX Sequence 149 BP; 54 A; 32 C; 12 G; 51 T; 0 U; 0 Other;
SQ
Query Match 94.3%; Score 26.4; DB 6; Length 149;
Best Local Similarity 96.4%; Pred. No. 1.6;
Matches 27; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GACATATTGTTATGTTGATGGAAC 28
DB 88 GACATATTGTTATGTTGATGGAAC 61

```

```

RESULT 44
AD37149
ID AD37149 standard; DNA; 459 BP.
XX
XX AD37149;
AC
XX 18-DEC-2003 (first entry)
DT
XX
XX 459bp HPV16 subgenomic fragment DNA sequence.
DE
XX
XX immortalised non-tumorigenic; human Schwann; schwannoma cell line;
KM screening cancer chemotherapeutic; antineoplastic activity;
KM carcinogenicity; neuroprotective; neurodegeneration; SV40;
KM Simian Virus 40; adenovirus; human papilloma virus; B6; E7; cytostatic;

```

KM gene therapy; cancer; HPV16; ds.
XX
OS Human papillomavirus type 16.
XX WO2003073996-A2.
XX
PD 12-SEP-2003.
XX
PF 03-MAR-2003; 2003WO-US006314.
XX
PR 01-MAR-2002; 2002US-0361528P.
XX
PA (HOU-S-) HOUSE EAR INST.
XX
PI Hung G, Li X;
XX
DR WPI; 2003-72206/68.
XX
PT Producing an immortalized non-tumorigenic human Schwann or schwannoma
PT cell line for treating cancer or neurodegenerative disorder comprises
PT selecting for immortalized cells that express the exogenous immortalizing
PT gene.
XX
XX
PS Disclosure; SEQ ID NO 1; 29pp; English.
XX
XX The invention relates to a the production of an immortalised non-
CC tumorigenic human Schwann or schwannoma cell line comprising: providing
CC a cell culture of human Schwann or schwannoma cells, introducing a
CC polynucleotide comprising an exogenous immortalising gene into the cells,
CC and selecting for immortalised cells that express the exogenous
CC immortalising gene and retain phenotypic properties of Schwann or
CC schwannoma cells. The invention further comprises: determining the effect
CC of a pharmacological agent on human Schwann or schwannoma cells;
CC screening cancer chemotherapeutic and antineoplastic activity of an agent
CC ; testing carcinogenicity of an agent; screening the neuroprotective
CC activity of an agent; treating neurodegeneration in a patient; and a kit
CC for screening a pharmacological agent on schwannomas cells. In the method
CC of producing an immortalised non-tumorigenic human Schwann or schwannoma
CC cell line, the polynucleotide is a subgenomic fragment of a virus
CC consisting of SV40 (Simian Virus 40), adenovirus and human papilloma
CC virus, comprising the B6 and E7 genes of the human papilloma virus
CC consisting of types 18, 31, 33 or 35, but preferably 16. The immortalised
CC non-tumorigenic human Schwann or schwannoma cell line has cytostatic and
CC neuroprotective activity and be useful in treating disorders by gene
CC therapy. The method is useful for producing an immortalised non-
CC tumorigenic human Schwann or schwannoma cell line useful for screening
CC cancer chemotherapeutic and antineoplastic activity of an agent, testing
CC carcinogenicity of an agent and screening a neuroprotective activity of
CC an agent and for preparing a composition for diagnosing or treating
CC neurodegenerative disorders and cancer. This polynucleotide represents
CC the DNA of a 459bp HPV16 sequence, a subgenomic fragment used in the
CC production of the immortalised non-tumorigenic human Schwann or
CC schwannoma cell line of the invention.
XX
SQ Sequence 459 BP; 157 A; 71 C; 100 G; 131 T; 0 U; 0 Other;
OY
Query Match 94.3%; Score 26.4; DB 10; Length 459;
Best Local Similarity 96.4%; Pred. No. 1.6;
Matches 27; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
DB 1 GACATTAATGTTAGTTGTATGGAAC 28
231 GATATATGTTATGTTATGTTATGGAAC 258
RESULT 45
AAL54445
ID AAL54445 standard; DNA; 477 BP.
XX
XX AAL54445;
XX
XX 03-APR-2003 (first entry)
XX

DE HPV16 E6 siRNA derived DNA sequence.
XX
XX Virucide; cytosstatic; anti-HIV; dermatological; small interfering RNA;
KM selective post-translational silencing; siRNA; oncogene; genital wart;
KM human papilloma virus; HPV gene; cancer; human cervical cancer; HIV;
KM smallpox; flu; common cold; cervical cancer; penile cancer;
KM malignant squamous cell carcinoma; verruca vulgaris; gene therapy; ds.
XX
XX Human papilloma virus.
XX
XX WO2003008573-A2.
XX
XX 30-JAN-2003.
XX
XX 17-JUL-2002; 2002WO-GB003300.
XX
XX 17-JUL-2001; 2001GB-00017358.
XX
XX 14-JAN-2002; 2002GB-00000688.
XX
XX 17-JUN-2002; 2002GB-00013855.
XX
XX (MILN/) MILNER A J.
XX
XX Milner AJ;
XX
XX WPI; 2003-221850/21.
XX
XX Selective post-transcriptional silencing of an exogenous viral gene (e.g.
PT human papilloma virus (HPV) E6), for treating e.g. cancer, comprises
PT using a small interfering RNA (siRNA) construct homologous to an mRNA of
PT the gene.
XX
XX Claim 11; Fig 11; 44pp; English.
XX
XX The invention relates to a novel method for selective post-translational
CC silencing in a mammalian cell of the expression of an exogenous gene of
CC viral origin. The method comprises introducing into the cell a small
CC interfering RNA (siRNA) construct that is homologous to a part of the
CC mRNA sequence of the gene. The method is useful for the selective post-
CC transcriptional silencing of an exogenous gene of viral origin (e.g. an
CC oncogene or human papilloma virus (HPV) gene) in a mammalian cell. The
CC method or the siRNA is particularly useful for treating cancer, human
CC cervical cancer, human immunodeficiency virus (HIV), smallpox, flu,
CC common cold, or a disease caused by a HPV (e.g. genital warts, cervical
CC cancer, penile cancer, malignant squamous cell carcinomas or verruca
CC vulgaris). An siRNA construct or vector is useful for use as a medicament
CC for the diseases mentioned. The polynucleotide sequence of the invention
CC can be used to treat disorders by gene therapy. This polynucleotide
CC sequence represents the DNA of a HPV siRNA sequence of the invention
XX
SQ Sequence 477 BP; 167 A; 75 C; 104 G; 131 T; 0 U; 0 Other;
OY
Query Match 94.3%; Score 26.4; DB 10; Length 477;
Best Local Similarity 96.4%; Pred. No. 1.6;
Matches 27; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
DB 1 GACATTAATGTTAGTTGTATGGAAC 28
251 GACATTAATGTTAGTTGTATGGAAC 278
RESULT 46
AEA51113
ID AEA51113 standard; DNA; 477 BP.
XX
XX AEA51113;
XX
XX 11-AUG-2005 (first entry)
XX
XX Human papillomavirus type 16 E6/E7 gene fragment, SEQ ID NO: 17.
XX
XX Delivery mechanism; gene therapy; cytosstatic; vulnery; virucide;
KM injury; infection; cancer; gastrointestinal disease;
KM gynecology and obstetrics; tumor; colorectal tumor;
KM

KM uterine cervix tumor; squamous cell carcinoma; neoplasm; E7 gene; ds;
XX B6 gene.
XX Human papillomavirus.
OS WO2005051431-A1.
XX
XX 09-JUN-2005.
PD
XX
PF 25-NOV-2004; 2004WO-GB004979.
XX
XX 25-NOV-2003; 2003GB-00027409.
PR 05-MAR-2004; 2004US-0549919P.
XX
XX (MILN/) MILNER A J.
PA
XX
XX Milner AJ;
PI
XX
XX WPI; 2005-405310/41.
DR
XX
XX Composition useful for delivering an agent into a cell comprises the
PT agent, a transfer agent and a solid or colloidal carrier medium.
XX
XX Disclosure; SEQ ID NO 17; 56pp; English.
PS
XX
XX The present invention relates to a method and composition comprising a
CC transfer agent and a solid or colloidal carrier medium for delivering a
CC biological agents into cells. The invention is useful for the treating
CC of cancer, tumor, carcinoma of cutaneous, squamous or cervical epithelia
CC and colorectal carcinoma, wounds, burns and scars. The invention is also
CC useful in gene therapy. The present sequence is the human papillomavirus
CC type 16 B6/E7 gene fragment. This sequence is useful in a method for
CC treating carcinomas.
CC
XX
SQ Sequence 477 BP; 167 A; 75 C; 104 G; 131 T; 0 U; 0 Other;
Query Match 94.3%; Score 26.4; DB 14; Length 477;
Best Local Similarity 96.4%; Pred. No. 1.6;
Matches 27; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GACATTATGTTATAGTTGTATGGAAC 28
DB 251 GACATTATGTTATAGTTGTATGGAAC 278
RESULT 47
ADP31984
ID ADP31984 standard; DNA; 543 BP.
XX
XX ADF31984;
AC
XX
XX 12-FEB-2004 (first entry)
DT
XX
XX Human papillomavirus fusion gene encoding sequence.
DE
XX
XX human papillomavirus; cervix cancer; ds.
KM
XX
XX Homo sapiens.
OS
XX
XX Key Location/Qualifiers
FH 1..543
FT CDS /*tag= a
FT
XX
XX CNI381583-A.
PN
XX
XX 27-NOV-2002.
PD
XX
XX 24-APR-2002; 2002CN-00117143.
PF
XX
XX 24-APR-2002; 2002CN-00117143.
PR
XX
XX (ONCO-) INST ONCOLOGY TUMOR HOSPITAL CHINESE ACA.
PA
XX

PI Zhao Q;
XX
XX WPI; 2003-258260/26.
DR
XX
XX P-PSDB; ADF31985.
DR
XX
XX Human papillomavirus B6/E7 fusion gene and its efficient expression
PT carrier and fusion protein vaccine.
XX
XX
XX Claim 3; SEQ ID NO 1; 16pp; Chinese.
PS
XX
XX The present invention relates to human papillomavirus B6/E7 fusion gene,
CC its preparing process, the process for configuring the efficient
CC expression carrier containing the gene and resultant expression carrier,
CC the fusion protein prepared from the gene, and the application of the
CC fusion gene and expression protein to medical science and medicine to
CC treat cervix cancer are disclosed. The present sequence represents the
CC human papillomavirus fusion gene encoding sequence.
CC
XX
SQ Sequence 543 BP; 185 A; 89 C; 118 G; 151 T; 0 U; 0 Other;
Query Match 94.3%; Score 26.4; DB 10; Length 543;
Best Local Similarity 96.4%; Pred. No. 1.7;
Matches 27; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GACATTATGTTATAGTTGTATGGAAC 28
DB 251 GACATTATGTTATAGTTGTATGGAAC 278
RESULT 48
ADO44067
ID ADO44067 standard; DNA; 747 BP.
XX
XX ADO44067;
AC
XX
XX 15-JUL-2004 (first entry)
DT
XX
XX Nucleotide sequence of an E7B6 fusion protein.
DE
XX
XX E6 protein; E7 protein; fusion protein; HPV16; HPV-associated cancer;
KM cervical cancer; immune response; lower gastrointestinal tract cancer;
KM anal cancer; reproductive system cancer; penile cancer; vulvar cancer;
KM gene; ss.
XX
XX Human papillomavirus type 16.
OS
XX
XX Synthetic.
XX
XX Key Location/Qualifiers
FH 1..747
FT CDS /*tag= a
FT /product= "E7B6 fusion protein"
FT
XX
XX WO2004030636-A2.
PN
XX
XX 15-APR-2004.
PD
XX
XX 02-OCT-2003; 2003WO-US031726.
PF
XX
XX 03-OCT-2002; 2002US-0415929P.
PR
XX
XX (AMHP) WYETH HOLDINGS CORP.
PA
XX
XX Smith L, Casasetti MC;
PI
XX
XX WPI; 2004-316328/29.
DR
XX
XX P-PSDB; ADO44066.
DR
XX
XX New polypeptide comprising human papillomavirus B6 and E7 polypeptides,
PT useful for treating or preventing human papillomavirus (HPV)-associated
XX cancers, e.g. cervical cancer.
XX
XX Example 1; Page 73; 101pp; English.
PS

CC The present sequence encodes an E7E6 fusion protein, comprising wild type
CC E7 and E6 polypeptides from human papillomavirus type 16 (HPV16). The
CC specification describes human papillomavirus E6 and E7 polypeptides,
CC where the E7 polypeptide has mutations at any one or more of the amino
CC acids corresponding to amino acids 24, 26 or 91 of the sequence given in
CC ADO44073 and the E6 polypeptide has no mutations or has mutations at any
CC one or more of the amino acids corresponding to amino acids 63 or 106 of
CC the sequence given in ADO44072. The polypeptides of the invention are
CC useful for treating or preventing human papillomavirus (HPV)-associated
CC cancers, such as cervical cancer. The fusion proteins and nucleic acids
CC encoding the fusion proteins are useful for generating immune responses
CC against HPV. They are also useful for treating lower gastrointestinal
CC tract cancer, e.g. anal cancer, and other cancers of the reproductive
CC system, including penile and vulvar cancer.

CC
XX
SQ Sequence 747 BP; 247 A; 132 C; 166 G; 202 T; 0 U; 0 Other;

Query Match 94.3%; Score 26.4; DB 12; Length 747;
Best Local Similarity 96.4%; Pred. No. 1.7;
Matches 27; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GACATTATTGTTATAGTTGTATGGAAC 28
Db 521 GACATTATTGTTATAGTTGTATGGAAC 548

RESULT 49
ADO44065
ID ADO44065 standard; DNA; 747 BP.
XX
AC ADO44065;
XX
DT 15-JUL-2004 (first entry)
XX
DE Nucleotide sequence of a fusion protein designated E6E7Pentm.
XX
XX B6 protein; E7 protein; fusion protein; HPV16; HPV-associated cancer;
KM cervical cancer; immune response; lower gastrointestinal tract cancer;
KM anal cancer; reproductive system cancer; penile cancer; vulvar cancer;
KM gene; ss.
XX
XX Human papillomavirus type 16.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT 1..747
FT CDS /*tag= a
FT /product= "E6E7Pentm fusion protein"
XX
PN MO2004030636-A2.
XX
PD 15-APR-2004.
XX
PF 02-OCT-2003; 2003WO-US031726.
XX
PR 03-OCT-2002; 2002US-0415929P.
XX
PA (AMHP) WYETH HOLDINGS CORP.
XX
PI Smith L, Cassecci MC;
XX
DR WPI: 2004-316328/29.
XX P-PSDB; ADO44064.
XX
XX New polypeptide comprising human papillomavirus E6 and E7 polypeptides,
PT useful for treating or preventing human papillomavirus (HPV)-associated
PT cancers, e.g. cervical cancer.
XX
XX
PS Claim 24; Page 71; 101pp; English.
XX
XX The present sequence encodes a fusion protein, comprising E6 and E7
CC polypeptides from human papillomavirus type 16 (HPV16). The fusion
CC protein is designated E6E7Pentm, and comprises an E6 amino terminus

CC (where residues 63 and 106 have been replaced with glycine) and an E7
CC carboxy terminus (where residues 24, 26 and 91 have been replaced with
CC glycine). E6E7Pentm is representative of fusion proteins of the
CC invention. The specification describes human papillomavirus E6 and E7
CC polypeptides, where the E7 polypeptide has mutations at any one or more
CC of the amino acids corresponding to amino acids 24, 26 or 91 of the
CC sequence given in ADO44073 and the E6 polypeptide has no mutations or has
CC mutations at any one or more of the amino acids corresponding to amino
CC acids 63 or 106 of the sequence given in ADO44072. The polypeptides of
CC the invention are useful for treating or preventing human papillomavirus
CC (HPV)-associated cancers, such as cervical cancer. The fusion proteins
CC and nucleic acids encoding the fusion proteins are useful for generating
CC immune responses against HPV. They are also useful for treating lower
CC gastrointestinal tract cancer, e.g. anal cancer, and other cancers of
CC the reproductive system, including penile and vulvar cancer.

CC
XX
SQ Sequence 747 BP; 246 A; 132 C; 171 G; 198 T; 0 U; 0 Other;

Query Match 94.3%; Score 26.4; DB 12; Length 747;
Best Local Similarity 96.4%; Pred. No. 1.7;
Matches 27; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GACATTATTGTTATAGTTGTATGGAAC 28
Db 230 GACATTATTGTTATAGTTGTATGGAAC 257

RESULT 50
ADO44061
ID ADO44061 standard; DNA; 747 BP.
XX
AC ADO44061;
XX
DT 15-JUL-2004 (first entry)
XX
DE Nucleotide sequence of an E6E7 fusion protein.
XX
XX B6 protein; E7 protein; fusion protein; HPV16; HPV-associated cancer;
KM cervical cancer; immune response; lower gastrointestinal tract cancer;
KM anal cancer; reproductive system cancer; penile cancer; vulvar cancer;
KM gene; ss.
XX
XX Human papillomavirus type 16.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT 1..747
FT CDS /*tag= a
FT /product= "E6E7 fusion protein"
XX
PN MO2004030636-A2.
XX
PD 15-APR-2004.
XX
PF 02-OCT-2003; 2003WO-US031726.
XX
PR 03-OCT-2002; 2002US-0415929P.
XX
PA (AMHP) WYETH HOLDINGS CORP.
XX
PI Smith L, Cassecci MC;
XX
DR WPI: 2004-316328/29.
XX P-PSDB; ADO44060.
XX
XX New polypeptide comprising human papillomavirus E6 and E7 polypeptides,
PT useful for treating or preventing human papillomavirus (HPV)-associated
PT cancers, e.g. cervical cancer.
XX
XX
PS Example 1; Page 68; 101pp; English.
XX
XX The present sequence encodes an E6E7 fusion protein, comprising wild type
CC E6 and E7 polypeptides from human papillomavirus type 16 (HPV16). The

CC specification describes human papillomavirus E6 and E7 polypeptides,
CC where the E7 polypeptide has mutations at any one or more of the amino
CC acids corresponding to amino acids 24, 26 or 91 of the sequence given in
CC ADO44073 and the E6 polypeptide has no mutations or has mutations at any
CC one or more of the amino acids corresponding to amino acids 63 or 106 of
CC the sequence given in ADO44072. The polypeptides of the invention are
CC useful for treating or preventing human papillomavirus (HPV)-associated
CC cancers, such as cervical cancer. The fusion proteins and nucleic acids
CC encoding the fusion proteins are useful for generating immune responses
CC against HPV. They are also useful for treating lower gastrointestinal
CC tract cancers, e.g. anal cancer, and other cancers of the reproductive
CC system, including penile and vulvar cancer.

XX
SQ Sequence 747 BP; 247 A; 132 C; 166 G; 202 T; 0 U; 0 Other;

Query Match 94.3%; Score 26.4; DB 12; Length 747;

Best Local Similarity 96.4%; Pred. No. 1.7;

Matches 27; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GACATTATGTGTAGTTGTATGGAAC 28

Db 230 GACATTATGTGTATGTTGTATGGAAC 257

Search completed: May 24, 2006, 05:55:53
Job time : 334.037 secs

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 24, 2006, 05:47:01 ; Search time 1932.77 Seconds
(without alignments)
810.102 Million cell updates/sec

Title: US-10-601-913-1
Perfect score: 28
Sequence: 1 GACATTATGTTATGATTGTGATGAGAAC 28

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 48236798 seqs, 27959665780 residues

Total number of hits satisfying chosen parameters: 96473596

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 1000 summaries

Database :

EST: *
1: gb_est1: *
2: gb_est3: *
3: gb_est4: *
4: gb_est5: *
5: gb_est6: *
6: gb_est7: *
7: gb_est2: *
8: gb_est7: *
9: gb_est8: *
10: gb_est9: *
11: gb_gsa1: *
12: gb_gsa2: *
13: gb_gsa3: *
14: gb_gsa4: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	26.4	94.3	244	3	B0106475
2	22.2	79.3	486	3	B0106475
3	21.6	77.1	499	11	AQ778171
4	21.2	75.7	658	13	CM302859
5	20.6	73.6	261	2	BM183005
6	20.6	73.6	324	2	BM531805
7	20.6	73.6	499	2	DN847201
8	20.6	73.6	832	14	DX977778
9	20.6	73.6	836	10	DT598349
10	20.6	73.6	848	9	CK897969
11	20.6	73.6	906	14	CR065011
12	20.6	72.9	1356	9	DN698710
13	20.4	72.9	677	14	AG111813
14	20.4	72.9	713	3	BM053630
15	20.4	72.9	724	3	BM206337
16	20.2	72.1	305	2	BF820181
17	20.2	72.1	351	9	DR165329
18	20.2	72.1	487	3	BO851812
19	20.2	72.1	543	14	DE110524

C 20	20.2	72.1	583	13	CL387597	CL387597
C 21	20.2	72.1	621	5	CD820121	CD820121
C 22	20.2	72.1	645	3	BO854577	BO854577
C 23	20.2	72.1	693	3	BO864783	BO864783
C 24	20.2	72.1	706	3	BO853474	BO853474
C 25	20.2	72.1	728	3	BO007386	BO007386
C 26	20.2	72.1	758	13	CM371871	CM371871
C 27	20.2	72.1	760	14	AG529207	AG529207
C 28	20	71.4	528	2	BO021561	BO021561
C 29	20	71.4	534	11	AQ829979	AQ829979
C 30	20	71.4	627	3	BM610169	BM610169
C 31	20	71.4	664	13	CZ821000	CZ821000
C 32	20	71.4	665	1	AL849178	AL849178
C 33	20	71.4	700	9	CK689149	CK689149
C 34	20	71.4	749	9	CK908094	CK908094
C 35	20	71.4	754	13	CZ754866	CZ754866
C 36	20	71.4	763	10	DV932465	DV932465
C 37	20	71.4	832	9	CK896534	CK896534
C 38	20	70.7	954	12	CG920252	CG920252
C 39	19.8	70.7	511	5	CF091813	CF091813
C 40	19.8	70.7	536	11	BZ392066	BZ392066
C 41	19.8	70.7	592	3	BO026745	BO026745
C 42	19.8	70.7	770	14	DX044921	DX044921
C 43	19.8	70.7	998	14	CNS07884	CNS07884
C 44	19.8	70.7	1268	14	AJ854876	AJ854876
C 45	19.6	70.0	170	5	CK928648	CK928648
C 46	19.6	70.0	232	2	BI034170	BI034170
C 47	19.6	70.0	297	10	DV017521	DV017521
C 48	19.6	70.0	299	5	CK553998	CK553998
C 49	19.6	70.0	305	2	BJ724011	BJ724011
C 50	19.6	70.0	340	3	BM857946	BM857946
C 51	19.6	70.0	351	10	R33750	R33750
C 52	19.6	70.0	371	11	AQ005319	AQ005319
C 53	19.6	70.0	400	5	CL187494	CL187494
C 54	19.6	70.0	405	14	CRU581204	CRU581204
C 55	19.6	70.0	423	11	B87841	B87841
C 56	19.6	70.0	430	7	BF221789	BF221789
C 57	19.6	70.0	430	7	BF589601	BF589601
C 58	19.6	70.0	436	12	CC090476	CC090476
C 59	19.6	70.0	456	11	BH273026	BH273026
C 60	19.6	70.0	461	13	CM229788	CM229788
C 61	19.6	70.0	467	14	DU0507309	DU0507309
C 62	19.6	70.0	486	5	CK766755	CK766755
C 63	19.6	70.0	493	13	CL579906	CL579906
C 64	19.6	70.0	494	13	CZ200459	CZ200459
C 65	19.6	70.0	519	9	DN303077	DN303077
C 66	19.6	70.0	524	9	DN290519	DN290519
C 67	19.6	70.0	552	13	CM881078	CM881078
C 68	19.6	70.0	558	9	DN293471	DN293471
C 69	19.6	70.0	571	4	BX564527	BX564527
C 70	19.6	70.0	573	10	DM126700	DM126700
C 71	19.6	70.0	579	2	BM029341	BM029341
C 72	19.6	70.0	580	10	DM552406	DM552406
C 73	19.6	70.0	601	13	CL583816	CL583816
C 74	19.6	70.0	608	10	DM082746	DM082746
C 75	19.6	70.0	613	5	CK763100	CK763100
C 76	19.6	70.0	615	5	CD304116	CD304116
C 77	19.6	70.0	615	10	DM141118	DM141118
C 78	19.6	70.0	617	10	DM141793	DM141793
C 79	19.6	70.0	618	5	CK551409	CK551409
C 80	19.6	70.0	622	3	BO995329	BO995329
C 81	19.6	70.0	623	9	DN290108	DN290108
C 82	19.6	70.0	624	3	BM319463	BM319463
C 83	19.6	70.0	633	12	CE817250	CE817250
C 84	19.6	70.0	634	3	BU006053	BU006053
C 85	19.6	70.0	635	13	CM151007	CM151007
C 86	19.6	70.0	636	1	AL883560	AL883560
C 87	19.6	70.0	637	1	AL893646	AL893646
C 88	19.6	70.0	639	3	BM372638	BM372638
C 89	19.6	70.0	645	3	BO991858	BO991858
C 90	19.6	70.0	653	10	DM081390	DM081390
C 91	19.6	70.0	654	3	BO985822	BO985822
C 92	19.6	70.0	655	10	DM143566	DM143566

969 19 67.9 289 2 BM159112 BM159112 fw27c12.y
970 19 67.9 289 2 BM181586 BM181586 fw33h04.y
971 19 67.9 289 2 BM532344 BM532344 fw06h11.y
972 19 67.9 289 3 BM811666 BM811666 fw09a08.y
973 19 67.9 289 3 BM811785 BM811785 fw10d12.y
974 19 67.9 289 3 BM077247 BM077247 fw14c11.y
975 19 67.9 289 3 BM077437 BM077437 fw16e06.y
976 19 67.9 289 3 BM077859 BM077859 fw18c11.y
977 19 67.9 289 2 BM840218 BM840218 fw18a09.y
978 19 67.9 289 2 BM266109 BM266109 fw39e12.y
979 19 67.9 289 2 BM530701 BM530701 fw27a01.y
980 19 67.9 289 3 BM075369 BM075369 fw24a08.y
981 19 67.9 289 3 BM077276 BM077276 fw24f07.y
982 19 67.9 289 3 BM077496 BM077496 fw17c02.y
983 19 67.9 289 3 BM077555 BM077555 fw17h05.y
984 19 67.9 289 3 BM077638 BM077638 fw18h06.y
985 19 67.9 289 3 BM078146 BM078146 fw85a11.y
986 19 67.9 289 3 BM078202 BM078202 fw85f12.y
987 19 67.9 289 3 BM480236 BM480236 fw87f04.y
988 19 67.9 289 3 BM480563 BM480563 fw89f04.y
989 19 67.9 289 3 BM618072 BM618072 fw88f08.y
990 19 67.9 289 3 BM618830 BM618830 fw80d03.y
991 19 67.9 289 2 BM157372 BM157372 fw41h06.y
992 19 67.9 289 2 BM159129 BM159129 fw27h10.y
993 19 67.9 289 2 BM315189 BM315189 fw33a06.y
994 19 67.9 289 3 BM076227 BM076227 fw22a11.y
995 19 67.9 289 3 BM076253 BM076253 fw22d04.y
996 19 67.9 289 3 BM076520 BM076520 fw88f07.y
997 19 67.9 289 3 BM076534 BM076534 fw89a05.y
998 19 67.9 289 3 BM076550 BM076550 fw89c02.y
999 19 67.9 289 3 BM076884 BM076884 fw20b10.y
1000 19 67.9 289 3 BM076889 BM076889 fw20c03.y

ALIGNMENTS

RESULT 1 244 bp mRNA linear EST 16-APR-2002
LOCUS B0106475
DEFINITION fc2102.e Rose Petals (Fragrant Cloud) Lambda Zap Express Library
ACCESSION B0106475
VERSION B0106475.1 GI:20156137
KEYWORDS EST.
SOURCE Rosa hybrid cultivar
ORGANISM Rosa hybrid cultivar
Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids I; Rosales; Rosaceae; Rosoideae; Rosa.
REFERENCE Guterman,I., Shalit,M., Menda,N., Pleskun,D., Dafny-Yelin,M.,
AUTHORS Shalev,G., Bar,E., Davydov,O., Ovadis,M., Emanuel,M., Wang,J.,
Adam,Z., Pichersky,E., Lewinson,E., Zamir,D., Vainstein,A. and
Weiss,D.
Rosa Scent: Genomics Approach to Discovering Novel Floral
Fragrance-Related Genes
JOURNAL Plant Cell 14 (10) 2325-2338 (2002)
COMMENT Contact: Naama Menda
Petal Genomics
Faculty of Agricultural, Food and Environmental Quality Sciences,
The Hebrew University of Jerusalem
P.O. Box 12, Rehovot, 76100, Israel
Tel: 972 8 9488 683
Fax: 972 8 9488 265
Email: shaham@agri.huji.ac.il
Seq primer: T3 forward.
Location/Qualifiers
FEATURES
source 1..244
/organism="Rosa hybrid cultivar"
/mol_type="mRNA"
/strain="Fragrant Cloud"
/db_xref="taxon:128735"

/clone="fc2102.e"
/issue_type="Petals"
/dev_stage="Young open flower at stage four"
/clone_lib="Rose Petals (Fragrant Cloud) Lambda Zap
Express Library"
/note="Vector: pBKCW, Site_1: EcoRI, Site_2: XhoI"
ORIGIN
Query Match 94.3%; Score 26.4; DB 3; Length 244;
Best Local Similarity 96.4%; Pred. No. 15;
Matches 27; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GACATTATGTTATGTTGATGGAAC 28
|||||
14 GACTTATGTTATGTTGATGGAAC 41

RESULT 2 486 bp mRNA linear EST 18-APR-2005
LOCUS DN832021/c
DEFINITION P026 Opisthorchis viverrini lambda Triplex2 Library Opisthorchis
viverrini cDNA 5', mRNA sequence.
ACCESSION DN832021
VERSION DN832021.1 GI:62727557
KEYWORDS EST.
SOURCE Opisthorchis viverrini
ORGANISM Opisthorchis viverrini
Eukaryote; Metazoa; Platyhelminthes; Trematoda; Digenea;
Opisthorchidae; Opisthorchiata; Opisthorchioidea; Opisthorchidae;
Opisthorchis.
REFERENCE 1 (bases 1 to 486)
AUTHORS Plaisor,P., Brindley,P.J., Laha,T., Loukas,A., Smout,M., Kaewkes,S.
and Srida,B.
Expressed sequence tags (ESTs) from cDNA libraries of adult stage
Opisthorchis viverrini
JOURNAL Unpublished (2005)
COMMENT Contact: Brindley, PJ
Department of Tropical Medicine
Tulane University, Health Sciences Center
1430 Tulane Avenue, New Orleans, LA 70112, USA
Tel: 1 504 988 4645
Fax: 1 504 988 6686
Email: paul.brindley@tulane.edu
Insert length: 486 Std Error: 0.00
Seq primer: 5' Triplex sequencing primer, CTCGAGACTGACGAGC
POLVA=yes.
Location/Qualifiers
FEATURES
source 1..486
/organism="Opisthorchis viverrini"
/mol_type="mRNA"
/strain="Khon Kaen Province, Thailand isolate"
/db_xref="taxon:6198"
/sex="hermaphrodite"
/tissue_type="whole body"
/dev_stage="adult"
/clone_lib="Opisthorchis viverrini lambda Triplex2
Library"

ORIGIN
Query Match 79.3%; Score 22.2; DB 9; Length 486;
Best Local Similarity 88.3%; Pred. No. 5.8e+02;
Matches 24; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 ACATTATGTTATGTTGATGGAAC 28
|||||
72 ACATTATGTTATGTTGATGGAAC 46

RESULT 3 499 bp DNA linear GSS 02-AUG-1999
LOCUS A0778171
DEFINITION HS_3206_A1_B05_MR CIT Approved Human Genomic Sperm Library D Homo
sapiens genomic clone Plate=3206 Col=9 Row=C, genomic survey

sequence.
 accession AQ78171 GI:5681131
 version AQ78171.1
 keywords GSS.
 source Homo sapiens (human)
 organism Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homidae; Homo.
 reference 1 (bases 1 to 499)
 authors Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T., Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and Hood,L.
 title Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome
 journal Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
 comment 10449764
 contact: Mahairas GG, Wallace JC, Hood L
 High Throughput Sequencing Center
 University of Washington
 401 Queen Anne Avenue North, Seattle, WA 98109, USA
 Tel: (206) 616-3618
 Fax: (206) 616-3887
 Email: jwallace@u.washington.edu
 Clones may be purchased from Research Genetics (info@resgen.com).
 BAC end Web Server: http://www.htsc.washington.edu
 plate: 3206 row: C column: 9
 seq primer: M13 Reverse
 class: BAC ends
 High quality sequence stop: 499.
 location/Qualifiers
 1..499
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /clone="plate=3206 Col=9 Row=C"
 /sex="male"
 /clone_1lb="CIT Approved Human Genomic Sperm Library D"
 /note="Organ: sperm; Vector: pBelbac11; BAC clones in E-Coli DH10B"
 ORIGIN
 Query Match 77.1%; Score 21.6; DB 11; Length 499;
 Best Local Similarity 85.7%; Pred. No. 9.7e+02;
 Matches 24; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 1 GACATTATGTTATGATTGTGATGGAAC 28
 Db 266 GAAACTACTCTTATGATTGTGATGGAAC 293
 RESULT 4
 CWS302859 658 bp DNA linear GSS 31-OCT-2004
 LOCUS 104 787 11464939 116 35676 078 Sorghum methylation filtered library
 DEFINITION (libID: 104) Sorghum bicolor genomic clone 11464939, genomic survey
 accession: CWS302859
 version CWS302859.1 GI:55019047
 keywords GSS.
 source Sorghum bicolor (sorghum)
 organism Sorghum bicolor
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae; Panicoideae; Andropogoneae; Sorghum.
 reference 1 (bases 1 to 658)
 authors Bedell,J.A., Budiman,M.A., Nunberg,A., Citek,R.W., Robbins,D., Jones,J., Flick,E., Rohlfing,T., Fries,J., Bradford,K., Korf,I.P., McNamey,J., Smith,M., Hoteman,H., Roe,B.A., Wiley,G., Korf,I.P., Rablinowicz,P.D., Lakey,N., McCombie,W.R., Jeddeloh,D.A. and Martensen,R.A.
 title Sorghum genome sequencing by methylation filtration
 journal Plos Biol. 3 (1), e13 (2005)

PUBMED 15660154
 comment Contact: Bedell JA
 Orion Genomics, LLC
 4041 Forest Park Ave, St. Louis, MO 63108, USA
 Tel: 314 615 6979
 Fax: 314 615 5975
 Email: jbedell@oriongenomics.com
 plate: 787 row: d column: 19
 seq primer: T3 Reverse
 class: methylation filtered
 High quality sequence stop: 658.
 location/Qualifiers
 1..658
 /organism="Sorghum bicolor"
 /mol_type="genomic DNA"
 /cultivar="Atx623"
 /db_xref="taxon:4558"
 /clone="11464939"
 /clone_1lb="Sorghum methylation filtered library (libID: 104)"
 /note="Organ: leaf; Vector: pBCSK(-); Site: 1; HindII; DNA prepared from purified nuclei was randomly sheared, end-repaired, size fractionated to enrich for the 0.5 to 5 kb fraction, ligated into HindII-digested pBCSK(-) vector and electroporated into E. coli cells. This is a methylation filtered library."
 ORIGIN
 Query Match 75.7%; Score 21.2; DB 13; Length 658;
 Best Local Similarity 88.5%; Pred. No. 1.4e+03;
 Matches 23; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 2 ACATTATGTTATGATTGTGATGGA 27
 Db 559 AATATTTTATGATTGTGCTCGAA 584
 RESULT 5
 BM183005 261 bp mRNA linear EST 11-DEC-2001
 LOCUS IMAGR:5566067 5', mRNA sequence.
 DEFINITION BM183005
 accession BM183005.1 GI:17513963
 version BM183005
 keywords EST.
 source Danio rerio (zebrafish)
 organism Danio rerio (zebrafish)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.
 reference 1 (bases 1 to 261)
 authors Clark,M., Johnson,S.L., Lehrach,H., Lee,R., Li,F., Marra,M., Eddy,S., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Pearson,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson,R.
 title Washu Zebrafish EST Project 1998
 journal Unpublished (1998)
 comment Contact: Stephen L. Johnson
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: zbratfish@wustl.edu
 title CDNA library construction by: Joe Barnes and Steve Johnson. DNA Sequencing by: Washington University Genome Sequencing Center Clone distribution: Research Genetics web address: http://www.researchgenetics.com/
 title Putative full length read
 The vector to vector length is 262
 zebrafish identity (p-value greater than 1e-99) found to: g[13461292]gb[A1105787]A1105787 db03d09 pl zF adult heart library

Danio rerio
Seq primer: T3 ET from Amerham.
Location/Qualifiers

1.261
/organism="Danio rerio"
/mol_type="mRNA"
/strain="C32"
/db_xref="taxon:7955"
/clone="IMAGE:556067"
/issue_type="embryo, 14 somite"
/lab_host="DH10B"
/clone_id="zebrafish C32 14 somite embryo"
/vector="PAMPI; Site 1: EcoRI; Site 2: NotI; First strand cDNA synthesis was primed using oligo-dt on magnetic beads with an additional primer
5'-ggcgccgctatagcactcacta-tagg-3'. Second strand synthesis was a 3-cycle PCR using the primers
5'-ggcgccgctatagcactcacta-3'
5'-aagcagtggtacacagcagagctt-tttttttttt-3'. cDNA was subsequently amplified in a 7-cycle PCR with the following primers: 5'-ggcgccgctatagcactcacta-3', 5'-aagcagtggt-acacagcag. Deoxy-UMP adaptors were added in a third PCR (5 cycles) and the primers
5'-caucaucaucaagcgctatagcactcacta-3' and
5'-caucaucaucaagcgctatagcactcacta-3'. Ends were treated with uracil DNA glycosylase and product with 3' overhangs was annealed to complementary ends of pAMP1. Insert can be excised using EcoRI and NotI. Library constructed by Joe Barnes and Steve Johnson (Washington University)."

ORIGIN

Query Match 73.6%; Score 20.6; DB 2; Length 261;
Best Local Similarity 85.2%; Pred. No. 2.3e+03;
Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 GACATTATTGTATAGTTGTATGAA 27
7 GACCTATTGTATAGTTATGTA 33

RESULT 6
BMS31805 324 bp mRNA linear EST 19-FEB-2002
LOCUS f91d07.v1 zebrafish sjd 2 day embryo Danio rerio cDNA clone
DEFINITION IMAGE:560652 5', mRNA sequence.
ACCESSION BMS31805
VERSION BMS31805.1 GI:18740694
KEYWORDS EST.
SOURCE Danio rerio (zebrafish)
ORGANISM Danio rerio
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.
1 (bases 1 to 324)
Clark, M., Johnson, S.L., Lehrach, H., Lee, R., Li, F., Marra, M., Rddy, S., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Stepcoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Page, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.
WashU Zebrafish EST Project 1998
Unpublished (1998)
Contact: Stephen L. Johnson
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: zbrfish@wustl.edu
cDNA library construction by: Joe Barnes and Steve Johnson. DNA sequencing by: Washington University Genome Sequencing Center Clone distribution: ReesourceZentrumPrimerBank, Berlin, Germany (web address: www.rzpd.de)

zebrafish identity (p-value greater than 1e-99) found to:
gt|3461292|gb|AI105787|AI105787 db03d09_pi zf adult heart library
Danio rerio
Seq primer: T3 ET from Amerham
High quality sequence stop: 324.
Location/Qualifiers

1.324
/organism="Danio rerio"
/mol_type="mRNA"
/strain="SJD"
/db_xref="taxon:7955"
/clone="IMAGE:5630652"
/issue_type="whole embryo, 2 day"
/lab_host="DH10B"
/clone_id="zebrafish sjd 2 day embryo"
/note="Vector: PAMPI; Site 1: EcoRI; Site 2: NotI; First strand cDNA synthesis was primed using oligo-dt on magnetic beads with an additional primer
5'-ggcgccgctatagcactcacta-tagg-3'. Second strand synthesis was a 3-cycle PCR using the primers
5'-ggcgccgctatagcactcacta-3' and
5'-aagcagtggtacacagcagagctt-tttttttttt-3'. cDNA was subsequently amplified in a 7-cycle PCR with the following primers: 5'-ggcgccgctatagcactcacta-3' and
5'-aagcagtggt-acacagcag. Deoxy-UMP adaptors were added in a third PCR (5 cycles) and the primers
5'-caucaucaucaagcgctatagcactcacta-3' and
5'-caucaucaucaagcgctatagcactcacta-3'. Ends were treated with uracil DNA glycosylase and product with 3' overhangs was annealed to complementary ends of pAMP1. Insert can be excised using EcoRI and NotI. Library constructed by Joe Barnes and Steve Johnson (Washington University)."

ORIGIN

Query Match 73.6%; Score 20.6; DB 2; Length 324;
Best Local Similarity 85.2%; Pred. No. 2.3e+03;
Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 GACATTATTGTATAGTTGTATGAA 27
7 GACCTATTGTATAGTTATGTA 33

RESULT 7
DN847201/c 499 bp mRNA linear EST 20-APR-2005
LOCUS KECB31-081.g-T7 NHH bovine early embryo cDNA library KECB Bos
DEFINITION taurus cDNA 3', mRNA sequence.
ACCESSION DN847201
VERSION DN847201.1 GI:62806361
KEYWORDS EST.
SOURCE Bos taurus (cattle)
ORGANISM Bos taurus
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Eumetazoa; Eumetazoa; Lauriatheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovinae; Bos.
1 (bases 1 to 499)
Olsaker, I. and Laerdahl, J.K.
Production of ESTs from a cDNA library of flushed bovine preimplantation embryos
Unpublished (2005)
Contact: Olsaker, I.
Dept. of Basic Sciences and Aquatic Medicine
Norwegian School of Veterinary Science
P.O. Box 8146 Dep, NO-0033 Oslo, Norway
Tel: +47 22964779
Fax: +47 22964758
Email: ingrid.olsaker@vet.hi.no
Single pass sequencing. Bases called and trimmed with Phred. Vector pCMV-PCR identified with crosses_match. EST name is Clone.g-T7
Seq primer: T7.
Location/Qualifiers

FEATURES

Tel: 925 296 5600
 Fax: 925 296 5710
 Email: cdna@jgi-psf.org
 Tissue Procurement: Timothy Grammer (Richard M. Harland Laboratory,
 University of California, Berkeley)
 http://tropicalis.berkeley.edu/home
 cDNA Library Preparation: DOE Joint Genome Institute:
 http://www.jgi.doe.gov
 DNA Sequencing: DOE Joint Genome Institute: http://www.jgi.doe.gov
 Clone Distribution: I.M.A.G.E. Consortium/LNL:
 http://image.lnl.gov
 Naming Conventions: EST name is generated by the concatenation of
 the JGI Clone id and the direction of sequencing. The suffix '.fwd'
 indicates a forward sequencing read of the insert. It does not
 necessarily reflect the orientation of the insert.
 plate: CAA0081 row: C column: 23
 High quality sequence stop: 767.
 Location/Qualifiers

FEATURES

source

1. 848
 /organism="Xenopus tropicalis"
 /mol_type="mRNA"
 /db_xref="taxon:8364"
 /clone="IMAGE:7681489"
 /sex="male"
 /tissue_type="Testes"
 /dev_stage="Adult"
 /lab_host="Electromax DH10B"
 /clone_lib="NIH XGC tropt63"
 /note="Vector: pCMVSPORT6; Site_1: SalI; Site_2: NotI;
 This library was made from dt primed cDNA and cloned into
 Invitrogen pCMVSPORT6 vector. The work was done at DOE
 Joint Genome Institute. Poly A RNA were primed with 5'
 GACTAGTTCAGATCGGAG CGAGCGCCCTTTTCTTTT 3'. cDNA
 were ligated to SalI adapter (5' TCGACCGACGCTCGG and
 5'CGGACGCTGGG), digested with NotI, size fractionated in
 1.1% agarose gel electrophoresis and ligated into NotI and
 SalI digested pCMVSPORT6 vector."

ORIGIN

Query Match 73.6%; Score 20.6; DB 9; Length 848;
 Best Local Similarity 85.2%; Pred. No. 2.3e+03;

Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 2 ACATTATTTGTTATGTTGATGGAAC 28

Db 244 ACAATTATTTGTTATGTTGATGGAAC 218

RESULT 11 906 bp DNA linear GSS 05-JUL-2004
 LOCUS CR065011
 DEFINITION Reverse strand read from insert in 5'HRPT insertion targeting and
 chromosome engineering clone MHPN42e12, genomic survey sequence.
 ACCESSION CR065011 GI:49798601
 VERSION GSS; genome survey sequence; MISCER.
 KEYWORDS Mus musculus (house mouse)
 SOURCE Mus musculus
 ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 906)
 Adams,D.J., Biggs,P.J., Cox,A.V., Davies,R.M., van der Weyden,L.,
 Jonkers,J., Smith,J., Plumb,R.W., Taylor,R.G., Nishijima,I., Yu,Y.,
 Rogers,J. and Bradley,A.

REFERENCE

AUTHORS

TITLE Direct Submission
 JOURNAL Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire,
 CB10 1SA, UK. http://www.sanger.ac.uk/MISCER

FEATURES

source

1. 906
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"

ORIGIN

/clone="MHPN42e12"
 /clone_lib="MHPN"

Query Match 73.6%; Score 20.6; DB 14; Length 906;
 Best Local Similarity 85.2%; Pred. No. 2.3e+03;
 Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 GACATTATTTGTTATGTTGATGGA 27

Db 572 GACATTATTTGTTATGTTGATGGA 598

RESULT 12 1356 bp mRNA linear EST 30-MAR-2005
 LOCUS DN698710/c

DEFINITION ClJ19-B11.3', mRNA sequence.

ACCESSION DN698710 GI:62060219

VERSION DN698710

KEYWORDS EST.

SOURCE Gasterosteus aculeatus (three spined stickleback)

ORGANISM Gasterosteus aculeatus

REFERENCE 1 (bases 1 to 1356)

AUTHORS Kingsley,D.M., Petchel,C., Balabhadra,S., Grimwood,J., Dickson,M.,
 Schmutz,J. and Myers,R.M.

TITLE Expressed sequence tags from Gasterosteus aculeatus

JOURNAL Unpublished (2003)

COMMENT Contact: Grimwood, Jane
 Stanford Human Genome Center
 Stanford University School of Medicine
 975 S California Ave, Palo Alto, CA 94304, USA
 Tel: 650 320 5917
 Fax: 650 320 5801
 Email: jane@shgc.stanford.edu
 plate: 19

High quality sequence start: 11
 High quality sequence stop: 899.
 Location/Qualifiers

FEATURES

source

1. 1356
 /organism="Gasterosteus aculeatus"
 /mol_type="mRNA"
 /strain="Bictrufjordur marine sticklebacks, Iceland"
 /db_xref="taxon:69293"
 /clone="ClJ19-B11"
 /sex="mixed male and female"
 /tissue_type="whole larva"
 /dev_stage="21 day old larvae collected at Svarup Stage 30
 (J Embryol. Exp. Morphol 6: 373-383,1958)"
 /lab_host="DH10B (T1 phage resistant)"
 /clone_lib="SHC-ClJ"
 /note="Vector: Express 1; Total and poly A+ RNA was
 isolated from the indicated stickleback tissue, and a cDNA
 library was constructed in the Express 1 plasmid vector by
 Open Biosystems. First strand cDNA synthesis was primed
 with an 54 bp linker primer containing an oligodT sequence
 preceded by a synthetic NotI site (first strand primer:
 5'-GACTAGTTCAGATCGGAGCGGCCGCTT(25-3'). Following
 second strand synthesis, cDNAs were made blunt at the end
 corresponding to the original 5 prime end of mRNA, and
 cloned directionally into the NotI and EcoRV sites of
 Express 1. Note that the EcoRV site is typically destroyed
 in the blunt end cloning, leaving a junction of the EcoRV
 'xxATC' (where is ATC is the second half of the EcoRV
 site, and xxx is derived from the cDNA sequence). A map of
 the Express 1 vector is available at:
 http://www.openbiosystems.com/cdna_library construction fa
 q.php# The primary library was transformed and amplified
 in DH10B (T1 phage resistant) bacteria. Clones available

from Open Biosystems:
http://www.openbiosystems.com/stickleback"

ORIGIN

Query Match 73.6%; Score 20.6; DB 9; Length 1356;
Best Local Similarity 85.2%; Pred. No. 2.3e+03;
Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 GACATTATTTGTTAAGTTGTATGGA 27

Db 35 GACATTATGTATATCATATGATGTAA 9

RESULT 13

AG111813 677 bp DNA linear GSS 03-NOV-2001
LOCUS Pan troglodytes DNA, clone: PTB-117P07.F, genomic survey sequence.
DEFINITION AG111813
ACCESSION AG111813
VERSION AG111813.1 GI:16732332
KEYWORDS GSS.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Pan.

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

Query Match 72.9%; Score 20.4; DB 3; Length 724;
 Best Local Similarity 95.5%; Pred. No. 2.8e+03;
 Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 ATTATGTTATAGTTGTATG 25
 |||||
 DB 508 ATTATGTTATAGTTGTATG 529

RESULT 16
 BE820181 205 bp mRNA linear EST 13-JAN-2001
 DEFINITION MR1-RT0025-171100-003-e09 RT0025 Homo sapiens cDNA, mRNA sequence.
 BE820181
 BE820181.1 GI:12158486

KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE
 AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
 Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
 Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H.,
 Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
 O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
 Simpson,A.U.

TITLE Shotgun sequencing of the human transcriptome with ORF expressed
 sequence tags

JOURNAL
 PUBMED Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 10737800

COMMENT Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br

FEATURES
 source This sequence was derived from the PAPSP/LICR Human Cancer Genome
 Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=MR1&c2=MR1-RT0025-
 171100-003-e09&c3=2000-11-17&c4=1)
 Seq primer: puc 18 forward
 High quality sequence start: 21
 High quality sequence stop: 94.
 Location/Qualifiers

1. 205
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /dev_stage="Adult"
 /clone_lib="RT0025"
 /note="Organ: kidney tumor; Vector: puc18; Site 1: SmaI;
 Site 2: SmaI; A mini-library was made by cloning products
 derived from ORESTS PCR (U.S. Letters Patent application
 No.196,716 - Ludwig Institute for Cancer Research)
 profiles into the pUC 18 vector. Reverse transcription of
 tissue mRNA and cDNA amplification were performed under
 low stringency conditions."

ORIGIN
 Query Match 72.1%; Score 20.2; DB 2; Length 205;
 Best Local Similarity 88.0%; Pred. No. 3.2e+03;
 Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 ACATATGTTATAGTTGTATGA 26
 |||||
 DB 90 ACATTTGTTTATAGTTCTATGA 114

RESULT 17
 DR165329 351 bp mRNA linear EST 20-JUN-2005
 LOCUS RRPPOS1.4_F02.b1_A029 Roots minus phosphorous Pinus taeda cDNA
 DEFINITION clone RRPPOS1_4_F02_A029 3', mRNA sequence.
 DR165329
 DR165329.1 GI:68023291

ACCESSION
 VERSION
 KEYWORDS EST.
 SOURCE Pinus taeda (loblolly pine)
 ORGANISM Pinus taeda

REFERENCE
 AUTHORS Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus.
 1 (bases 1 to 351)
 Pratt,L., Cordonnier-Pratt,M.M., Lorenz,W.W., Zimmermann,C. and
 Dean,J.F.D.

TITLE An EST database from phosphorous-deficient loblolly pine (Pinus
 taeda) roots
 JOURNAL Unpublished (2005)
 COMMENT Other ESTs: RRPPOS1_4_F02.g1_A029
 Contact: Cordonnier-Pratt MM
 Laboratory for Genomics and Bioinformatics
 The University of Georgia, Department of Plant Biology
 Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
 Tel: 706 542 1860
 Fax: 706 583 0210
 Email: mmpcr@uga.edu

RNA prepared and library constructed by W. Walter Lorenz (School of
 Forest Resources, University of Georgia); plant material prepared
 by Craig Zimmermann (School of Forest Resources, University of
 Georgia) using rooted cuttings provided by the Forest Biology
 Research Cooperative (FBRC) and the CCLONBS project a the
 University of Florida; sequencing done in the Laboratory for
 Genomics and Bioinformatics, University of Georgia. Sequence ends
 have been trimmed to exclude vector and regions below Phred quality
 16. Three-prime sequences are presented as their reverse complement
 and have been trimmed to exclude polyA.
 Seq primer: M13-21 (TGTAAACGACGCGCCAGT)
 POLYA=yes.
 Location/Qualifiers

1. 351
 /organism="Pinus taeda"
 /mol_type="mRNA"
 /strain="3 CCLONBS"
 /db_xref="taxon:3352"
 /clone="RRPOS1_4_F02_A029"
 /lab_host="DH10B-T1 pFage-resistant E. coli"
 /clone_lib="Roots minus phosphorous"
 /note="Organ: Root; Vector: pSL1180; Site_1: EcoRI;
 Site_2: XhoI; The library was prepared from polyA+ RNA
 from the roots of 1-year-old loblolly pine (Pinus taeda)
 cuttings that were rooted and then planted in washed sand.
 The rooted cuttings were maintained for 139 days (July 28
 2003 harvest) under ambient conditions in a local
 greenhouse. They were kept on a weekly regimen of 0.5x
 nutrient-complete Hoagland's solution and supplemented
 with additional water sufficient to maintain a 15% soil
 moisture content. Starting five weeks prior to harvesting
 roots for RNA preparation, the weekly fertilizer regimen
 for the potted trees was modified to omit phosphorus from
 the 0.5x Hoagland's solution (5 applications). For two
 days prior to starting the modified fertilizer regimen,
 pots were flushed extensively with water to reduce
 residual levels of phosphorus. Double-stranded cDNA was
 cloned unidirectionally into pSL1180. Inserts can be
 excised with EcoRI (5' end) and XhoI (3' end)."

ORIGIN
 Query Match 72.1%; Score 20.2; DB 9; Length 351;
 Best Local Similarity 88.0%; Pred. No. 3.2e+03;
 Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 ACATATGTTATAGTTGTATGA 26
 |||||

Db 83 ACATTATTGTTATGATTGGACAGA 107

RESULT 18
BQ851812
LOCUS
DEFINITION OGB16F20.1y.ab1 OG ABCDI lettuce salinas Lactuca sativa cDNA clone
OGB16F20. mRNA sequence.

ACCESSION BQ851812
VERSION BQ851812.1 GI:22237277

KEYWORDS
SOURCE Lactuca sativa
ORGANISM Lactuca sativa

REFERENCE
AUTHORS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; asterids; campanulids; Asterales; Asteraceae; Cichorioideae; Cichorieae; Lactuca.
1 (bases 1 to 487)
Kozik, A., Michelmore, R.W., Knapp, S., Matvienko, M., Rieseberg, L., Lin, H., van Damme, M., Lavelle, D., Chevalier, P., Ziegler, J., Ellison, P., Kolkmann, J., Slabaugh, M.S., Livingston, K., Zhou, Y., Lai, Z., Church, S., Jackson, L. and Bradford, K.
Lettuce and Sunflower ESTs from the Compositae Genome Project
http://compomics.ucdavis.edu/
Unpublished (2002)

JOURNAL
COMMENT Contact: Alexander Kozik [R.W.Michelmore]
Department of Vegetable Crops, R.W.Michelmore Lab
University of California at Davis (UCD)
Aemundson Hall, UCD, Davis, CA 95616, USA
Tel: 1-(530)-742-1742
Fax: 1-(530)-752-9659
Email: akozik@ucdavis.edu [michelmore@vegmail.ucdavis.edu]
belongs to contig OG_Ca_contig6733, see http://cgdb.ucdavis.edu/
for details.
Plate: OGB16 row: F column: 20.

FEATURES
source
1..487
Location/Qualifiers
/organism="Lactuca sativa"
/mol_type="mRNA"
/cultivar="Salinas"
/db_xref="taxon:4236"
/clone="OGB16F20"
/lab_host="E.coli"
/clone_lib="OG_ABCDI lettuce salinas"
/note="Vector: pBRCDNA51AB. The library was constructed from 10 different sources of RNA from a single genotype. Separate cDNAs were generated using primers that incorporated unique 5' and 3' tags to distinguish each source of RNA. cDNAs were then pooled, size-fractionated, directionally cloned into a custom medium-copy vector and transformations made with four size classes to minimize size bias. Details of each source of RNA and library construction can be obtained at http://cgdb.ucdavis.edu/
TAG LIB=OG ABCDI lettuce salinas
TAG_SEQ=TTTACCGGG"

ORIGIN

Query Match 72.1%; Score 20.2; DB 3; Length 487;
Best Local Similarity 88.0%; Pred. No. 3.2e+03;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GACATTATTGTTATGATTGGATG 25
|||||
Db 102 GACATTCTGTATGATTGGATG 126
|||||

RESULT 19
DE110524
LOCUS
DEFINITION Oryzias latipes DNA, reverse end of BAC clone: Md0153B05, genomic survey sequence.

ACCESSION DE110524

VERSION DE110524.1 GI:71610277
GSS.
KEYWORDS Oryzias latipes (Japanese medaka)
SOURCE Oryzias latipes
ORGANISM Oryzias latipes

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha; Belontiiformes; Adrianichthyidae; Oryziinae; Oryzias.
1
Shimizu, N., Asakawa, S., Shimizu, A. and Sasaki, T.
The BAC end sequence of Oryzias latipes
Published Only in Database (2005)

JOURNAL
COMMENT 2 (bases 1 to 543)
Shimizu, N., Asakawa, S., Sasaki, T. and Shimizu, A.
Direct Submission
Submitted (30-JUN-2005) Nobuyoshi Shimizu, Keio University, School of Medicine, Molecular Biology, 35 Shinanomachi, Shinjuku-Ku, Tokyo, 160-8582, Japan (E-mail: nshimizu@med.keio.ac.jp, Tel: 81-3-3351-2370, Fax: 81-3-3351-2370)

FEATURES
source
1..543
Location/Qualifiers
/organism="Oryzias latipes"
/mol_type="genomic DNA"
/db_xref="taxon:8090"
/clone="Md0153B05"
/clone_lib="Medaka Hdr BAC library"
/note="This sequence is reverse end of BAC clone Md0153B05."

ORIGIN

Query Match 72.1%; Score 20.2; DB 14; Length 543;
Best Local Similarity 88.0%; Pred. No. 3.3e+03;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 4 ATTATGTTATGATTGATGAAC 28
|||||
Db 272 ATTATGTTATGATTGATGAAC 296
|||||

RESULT 20
CL387597/c 583 bp DNA linear GSS 19-AUG-2004
LOCUS RRC144.281113.r RPI-44 Sus scrofa genomic clone RRC144_281113.
DEFINITION Genomic survey sequence.

ACCESSION CL387597
VERSION CL387597.1 GI:51439557
GSS.
SOURCE Sus scrofa (pig)
ORGANISM Sus scrofa

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Suina; Suidae; Sus.
1 (bases 1 to 583)
Rogacheva, M.B., Meyers, S., He, W., Larkin, D.M., Marron, B.M., Beaver, J.B. and Schook, L.B.
Piggy-BACing the Human Genome: Constructing a Porcine Physical Map Through Comparative Genomics
Unpublished (2004)

JOURNAL
COMMENT Other GSSs: RRC144.281113.f
Contact: Lawrence B. Schook
Department of Animal Sciences
University of Illinois at Urbana Champaign
1201 W. Gregory Dr., Urbana, IL 61801, USA
Tel: 217 265 5326
Fax: 217 244 5617
Email: schook@uiuc.edu
Clones are derived from the porcine BAC library RPI-44
(http://www.bacpac.chori.org/porcine242.htm). For BAC library availability, please contact Pieter de Jong (pdjong@chori.org).
Clones may be purchased from BACPAC Resources
(http://BACPACResources.chori.org). This work was undertaken as part of the International Swine Genome Sequencing Consortium by University of Illinois at Urbana Champaign, USA with funds provided

by Grant No. AG2002-34480-11828 from USDA-CSREES and
AG2001-35205-09965 from USDA/NRI (Livestock Genome Sequencing
Initiative)

Plate: 281 row: L column: 13
Seq primer: SP6
Class: BAC ends
Location/Qualifiers

FEATURES
source

1..583
/organism="Sus scrofa"
/mol_type="genomic DNA"
/strain="four pigs (bred: 37.5% Yorks Landrace and 25%
Meishan)"
/db_xref="taxon:9823"
/clone="RPC14_281L13"
/sex="male"
/cell_type="Blood"
/clone_lib="RPC1-44"
/note="Vector: pFARAC2; Site_1: EcoRI; Site_2: EcoRI;
porcine male BAC library produced by Pieter de Jong"

ORIGIN

Query Match 72.1%; Score 20.2; DB 13; Length 583;
Best Local Similarity 88.0%; Pred. No. 3.3e+03;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 CATTATGTTATGTTGATGGA 27
|||||
Db 459 CATTATGTTATGTTTGGTA 435

RESULT 21
CD820121 621 bp mRNA linear EST 10-JUL-2003
LOCUS BN20.051E05F011220 BN20 Brassica napus cDNA clone BN20051B05, mRNA
DEFINITION
Sequence.
ACCESSION CD820121 GI:32502061
VERSION CD820121.1
KEYWORDS EST.
SOURCE Brassica napus (rape)
ORGANISM Brassica napus
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
1 (bases 1 to 621)
Genoplante.
Genoplante, a major partnership french program in plant genomics
unpublished (2003)
Contact: Genoplante
Genoplante
93, rue Henri Rochefort 91025 EVRY CEDEX France
Tel: 33 1 69 47 54 00
Fax: 33 1 69 47 54 10
This sequence has been generated in the framework of the french
plant genomics programme 'Genoplante' (<http://www.genoplante.com>
and <http://genoplante-info.infobiogen.fr>).
Location/Qualifiers

FEATURES
source

1..621
/organism="Brassica napus"
/mol_type="mRNA"
/cultivar="Jec Neuf"
/db_xref="taxon:3708"
/clone="BN20051B05"
/issue_type="seed"
/clone_lib="BN20"

ORIGIN

Query Match 72.1%; Score 20.2; DB 5; Length 621;
Best Local Similarity 88.0%; Pred. No. 3.3e+03;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 ACATTATGTTATGTTGATGGA 26
|||||
Db 517 AGATTATGTTATGTTGATGGA 541

RESULT 22
B0854577 645 bp mRNA linear EST 14-AUG-2002
LOCUS OGB23K09.y5.ab1 OG ABCDI lettuce salinas Lactuca sativa cDNA clone
DEFINITION OGB23K09, mRNA sequence.
B0854577
B0854577.1 GI:22240042
EST.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Lactuca sativa
Lactuca sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
asterids; campanulids; Asterales; Asteraceae; Cichorioideae;
Cichorieae; Lactuca.
1 (bases 1 to 645)

REFERENCE

AUTHORS

Kozik, A., Michelmore, R.W., Knapp, S., Matvienko, M., Rieseberg, L.,
Lin, H., van Damme, M., Lavelle, D., Chevalier, P., Ziegler, J.,
Rilison, P., Kolkmann, J., Slabaugh, M.S., Livingston, K., Zhou, Y.,
Lai, Z., Church, S., Jackson, L. and Bradford, K.
Lettuce and Sunflower ESTs from the Compositeae Genome Project
<http://comgenomics.ucdavis.edu/>
Unpublished (2002)
Contact: Alexander Kozik [R.W.Michelmore]
Department of Vegetable Crops, R.W.Michelmore Lab
University of California at Davis (UCD)
Amundson Hall, UCD, Davis, CA 95616, USA
Tel: 1-(530)-742-1742
Fax: 1-(530)-752-9659
Email: akozik@ucdavis.edu [michelmore@vegmail.ucdavis.edu]
belongs to config OG_CA_Config6733, see <http://cspdb.ucdavis.edu/>
for details.
Plate: OGB23 row: K column: 09.

JOURNAL
COMMENT

FEATURES
source

1..645
/organism="Lactuca sativa"
/mol_type="mRNA"
/cultivar="Salinas"
/db_xref="taxon:4236"
/clone="OGB23K09"
/lab_host="E.coli"
/clone_lib="OG ABCDI lettuce salinas"
/note="Vector: pBRCDNA5flAB; The library was constructed
from 10 different sources of RNA from a single genotype.
Separate cDNAs were generated using primers that
incorporated unique 5' and 3' tags to distinguish each
source of RNA. cDNAs were then pooled, size-fractionated,
directionally cloned into a custom medium-copy vector and
transformations made with four size classes to minimize
size bias. Details of each source of RNA and library
construction can be obtained at <http://cspdb.ucdavis.edu/>
TAG_LTB-OG ABCDI lettuce salinas
TAG_SEQ=GTGAGCGGG"

ORIGIN

Query Match 72.1%; Score 20.2; DB 3; Length 645;
Best Local Similarity 88.0%; Pred. No. 3.3e+03;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GACATTATGTTATGTTGATG 25
|||||
Db 102 GACATTCTGTTATGTTGATG 126

RESULT 23

B0864783

LOCUS

DEFINITION OGC27K11.y5.ab1 OG ABCDI lettuce salinas Lactuca sativa cDNA clone
OGC27K11, mRNA sequence.
B0864783
B0864783.1 GI:22250248

ACCESSION
VERSION

KEYWORDS	EST.
SOURCE	Lactuca sativa
ORGANISM	Lactuca sativa
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; asterids; campanulids; Asterales; Asteraceae; Cichorioideae; Cichoriaceae; Lactuca.
AUTHORS	1 (bases 1 to 693) Kozik,A., Michelmore,R.W., Knapp,S., Matvienko,M., Rieseberg,L., Lin,H., van Damme,M., Lavelle,D., Chevalier,P., Ziegler,U., Ellison,P., Kolman,J., Slabaugh,M.S., Livingston,K., Zhou,Y., Lai,Z., Church,S., Jackson,L. and Bradford,K.
TITLE	Lettuce and Sunflower ESTs from the Compositae Genome Project http://compgenomics.ucdavis.edu/unpublished (2002)
JOURNAL	Contact: Alexander Kozik [R.W.Michelmore] Department of Vegetable Crops, R.W.Michelmore Lab University of California at Davis (UCD) Aamundson Hall, UCD, Davis, CA 95616, USA Tel: 1-(530)-742-1742 Fax: 1-(530)-752-9659 Email: akozik@ucdavis.org [michelmore@vegmail.ucdavis.edu] belongs to contig OG_CA_Config6733, see http://cgdb.ucdavis.edu/ for details.
FEATURES	Plate: QGC27 row: K column: 11. Location/Qualifiers 1..693 /organism="Lactuca sativa" /mol_type="mRNA" /cultivar="Salinas" /db_xref="taxon:4236" /clone="QGC27K11" /lab_host="E.coli" /clone_lib="QG ABCDI lettuce salinas" /note="Vector: pBRCDNA5f1ab. The library was constructed from 10 different sources of RNA from a single genotype. Separate cDNAs were generated using primers that incorporated unique 5' and 3' tags to distinguish each source of RNA. cDNAs were then pooled, size-fractionated, directionally cloned into a custom medium-copy vector and transformations made with four size classes to minimize size bias. Details of each source of RNA and library construction can be obtained at http://cgdb.ucdavis.edu/TAG_LIB=QG_ABCDI lettuce salinas TAG_SEQ=TGAGCCGGG9"
ORIGIN	
Query Match	72.1%; Score 20.2; DB 3; Length 693;
Best Local Similarity	88.0%; Pred. No. 3.3e+03;
Matches	22; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
QY	1 GACATTATGTTAGTTGATG 25
Db	102 GACATTCGTGTTGATGATG 126
RESULT 24	
LOCUS	B0853474 706 bp mRNA linear EST 14-AUG-2002
DEFINITION	QGC20020.yg.ab1 QG ABCDI lettuce salinas Lactuca sativa cDNA clone
ACCESSION	QGC20020, mRNA sequence.
VERSION	B0853474
KEYWORDS	B0853474.1 GI:22238939
SOURCE	EST.
ORGANISM	Lactuca sativa
ORGANISM	Lactuca sativa
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; asterids; campanulids; Asterales; Asteraceae; Cichorioideae; Cichoriaceae; Lactuca.
AUTHORS	1 (bases 1 to 706) Kozik,A., Michelmore,R.W., Knapp,S., Matvienko,M., Rieseberg,L.,

TITLE	JOURNAL	COMMENT
Lin,H., van Damme,M., Lavie,D., Chevalier,P., Ziegler,J., Ellison,P., Kolkmann,J., Slabaugh,M.S., Livingston,K., Zhou,Y., Lai,Z., Church,S., Jackson,L. and Bradford,K. Lettuce and Sunflower ESTs from the Compositae Genome Project http://compgenome.ucdavis.edu/ Unpublished (2002)	Department of Vegetable Crops, R.W.Mitchellmore Lab University of California at Davis (UCD) Aamundson Hall, UCD, Davis, CA 95616, USA Tel: 1-(530)-742-1742 Fax: 1-(530)-752-9659 Email: akozik@ucdavis.edu [mitchellmore@vegmail.ucdavis.edu] belongs to contig OG_Ca_contig6733, see http://cgdb.ucdavis.edu/ for details.	
FEATURES	source	
1..706	location/Qualifiers	
/organism="Lactuca sativa"		
/mol_type="mRNA"		
/cultivar="Salinas"		
/db_xref="taxon:4236"		
/clone="QGB20J20"		
/lab_host="E.coli"		
/clone_lib="OG_ABCDI lettuce salinas"		
/note="Vector: pBRCDNA51ab. The library was constructed from 10 different sources of RNA from a single genotype. Separate cDNAs were generated using primers that incorporated unique 5' and 3' tags to distinguish each source of RNA. cDNAs were then pooled, size-fractionated, directionally cloned into a custom medium-copy vector and transformations made with four size classes to minimize size bias. Details of each source of RNA and library construction can be obtained at http://cgdb.ucdavis.edu/ TAG TISSUE=chemical induction TAG LIB=OG ABCDI lettuce salinas TAG_SEQ=TTGATCCGGG"		
ORIGIN		
Query Match	72.1%; Score 20.2; DB 3; Length 706;	
Best Local Similarity	88.0%; Pred. No. 3.3e+03;	
Matches	22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;	
Qy	1 GACATTATGTTATGATGTTGTATG 25	
Db	102 GACATTCCTGTTATGTTGTATGATG 126	
RESULT 25		
LOCUS	BU007386	
DEFINITION	OGH2a10.y5.ab1 OG EFGHJ lettuce serriola Lactuca serriola cDNA	
ACCESSION	BU007386	
VERSION	BU007386.1	
KEYWORDS	GI:22441781	
SOURCE	EST.	
ORGANISM	Lactuca serriola	
	Lactuca serriola	
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; asterids; campanulids; Asterales; Asteraceae; Cichorioideae; Cichorieae; Lactuca.	
REFERENCE	1 (bases 1 to 728)	
AUTHORS	Kozik,A., Mitchellmore,R.W., Knapp,S., Matvienko,M., Rieseberg,L., Lin H., van Damme,M., Lavelle,D., Chevalier,P., Ziegler,J., Ellison,P., Kolkmann,J., Slabaugh,M.S., Livingston,K., Zhou,Y., Lai,Z., Church,S., Jackson,L. and Bradford,K. Lettuce and Sunflower ESTs from the Compositae Genome Project http://compgenome.ucdavis.edu/ Unpublished (2002)	
JOURNAL	COMMENT	
CONTACT: Alexander Kozik [R.W.Mitchellmore]	Department of Vegetable Crops, R.W.Mitchellmore Lab	
University of California at Davis (UCD)		

ORIGIN
 /clone="MSG01-438K16.T7"
 /sex="male"
 /tissue type="mixture of kidney and spleen"
 /clone_id="MSG01 Mouse Male BAC Library"

Query Match 72.1%; Score 20.2; DB 14; Length 760;
 Best Local Similarity 88.0%; Pred. No. 3.3e+03;
 Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GACATTATGTTATAGTTGTATG 25
 |||||
 Db 289 GACATTATGTTATAGTTGTATG 265

RESULT 28
 BG021561/c 528 bp mRNA linear EST 19-FEB-2003
 LOCUS dg32b09.x1 Xenopus laevis gastrula non normalized Xenopus laevis
 DEFINITION cDNA clone IMAGE:3750473 3', mRNA sequence.

ACCESSION BG021561
 VERSION BG021561.1 GI:12477640
 KEYWORDS EST.
 SOURCE Xenopus laevis (African clawed frog)
 ORGANISM Xenopus laevis

REFERENCE
 AUTHORS Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Amphibia; Batrachia; Anura; Mesodactylia; Pipridae; Pipridae;
 Xenopodinae; Xenopus; Xenopus.
 1 (bases 1 to 528)
 Clifton, S., Johnson, S.L., Blumberg, B., Song, J., Miller, L.,
 Pape, D., Martin, J., Wylie, T., Underwood, K., Theising, B., Bowers, Y.,
 Peterson, R., Gibbons, M., Harvey, N., Ritter, E., Jackson, Y., McCann, R.,
 Watsen, R. and Wilson, R.
 Washu Xenopus EST Project, 1999
 Unpublished (1999)

TITLE JOURNAL
 COMMENT
 Contact: Sandy Clifton, Ph.D.
 Washu Xenopus EST Project, 1999
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watsen.wustl.edu

Library constructed by Bruce Blumberg
 Library normalized by Jihwan Song
 DNA Sequencing by: Washington University Genome Sequencing Center
 Source lab clone id - xlmga009c18 This clone is available
 royalty-free through LML; contact the IMAGE Consortium
 (image@image.llnl.gov) for further information.
 Seq primer: -40up from Gibco
 High quality sequence stop: 476.
 Location/Qualifiers

1. 528
 Location/Qualifiers

FEATURES
 SOURCE
 /organism="Xenopus laevis"
 /mol_type="mRNA"
 /db_xref="taxon:8355"
 /clone="IMAGE:3750473"
 /tissue type="gastrula (stages 10.5, 11.5 mixed)"
 /lab host="Top-10 P"
 /clone_id="Xenopus laevis gastrula non normalized"
 /note="Vector: Bluescript SK-; site 1: EcoRI; site 2:
 XhoI; cDNA was prepared from 2 ug of poly A+ RNA (equal
 parts from stage 10.5 and stage 11.5 gastrulae).
 EcoRI-XhoI cut cDNA was then ligated into UniZap-XR
 (Stratagene) with EcoRI at the 5' end and XhoI at the 3'
 end. The library was mass excised and used to infect
 Top10P. Clones were picked into freeing medium (per
 kit) 15 g tryptone, 10g yeast extract, 5g NaCl, 36 mM
 KH2PO4, 13.2 mM KH2PO4, 1.7 mM Na-citrate, 0.4 mM MgSO4 7
 H2O, 6.8 mM (NH4)2SO4, 4 % w/v glycerol and grown for 24
 hours. Original library construction by Bruce Blumberg
 (Cho et al 1991 Cell 67, 1111-1120)."

ORIGIN

Query Match 71.4%; Score 20; DB 2; Length 528;
 Best Local Similarity 82.1%; Pred. No. 3.9e+03;
 Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 GACATTATGTTATAGTTGTATG 28
 |||||
 Db 44 GACATTATGTTATAGTTGTATG 17

RESULT 29
 A0829979/c 534 bp DNA linear GSS 27-AUG-1999
 LOCUS HS 4828 B1 F06 T7A CIT Approved Human Genomic Sperm Library D Homo
 DEFINITION sapiens genomic clone Plate=4828 Col=11 Row=L, genomic survey
 sequence.

ACCESSION A0829979
 VERSION A0829979.1 GI:5796041
 KEYWORDS GSS.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE
 AUTHORS Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homnidae; Homo.
 1 (bases 1 to 534)
 Mahairas, G.G., Wallace, J.C., Smith, K., Swartzell, S., Holzman, T.,
 Keller, A., Shaker, R., Furlong, J., Young, J., Zhao, S., Adams, M.D. and
 Hood, L.
 Sequence-tagged connectors: A sequence approach to mapping and
 scanning the human genome
 Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
 10449764
 Contact: Mahairas GG, Wallace JC, Hood L
 High Throughput Sequencing Center
 University of Washington
 401 Queen Anne Avenue North, Seattle, WA 98109, USA
 Tel: (206) 616-3618
 Fax: (206) 616-3887
 Email: jwallace@u.washington.edu
 Clones may be purchased from Research Genetics (info@resgen.com).
 BAC end Web Server: http://www.hnsc.washington.edu
 Plate: 4828 row: L column: 11
 Seq primer: T7
 Class: BAC ends
 High quality sequence stop: 534.
 Location/Qualifiers

1. 534
 Location/Qualifiers

FEATURES
 SOURCE
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /clone="plate=4828 Col=11 Row=L"
 /sex="male"
 /clone_id="CIT Approved Human Genomic Sperm Library D"
 /note="Organ: sperm; Vector: pBelBAC11; BAC Clones in
 E-Coli DH10B"

ORIGIN

Query Match 71.4%; Score 20; DB 11; Length 534;
 Best Local Similarity 100.0%; Pred. No. 3.9e+03;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 ATTATGTTATAGTTGTAT 23
 |||||
 Db 365 ATTATGTTATAGTTGTAT 346

RESULT 30
 BM610169/c 627 bp mRNA linear EST 25-FEB-2002
 LOCUS 17006871.10115 A.Gam ad.cDNA.blood1 Anopheles gambiae cDNA clone
 DEFINITION 19600449718565 5', mRNA sequence.

ACCESSION BM610169
 VERSION BM610169.1 GI:18908273
 KEYWORDS EST.

```

SOURCE      Anopheles gambiae (African malaria mosquito)
ORGANISM    Anopheles gambiae
REFERENCE   1 (bases 1 to 627)
AUTHORS     Holt,R.A., Lin,J.-O., Murphy,S.D., Evans,C.A., Kraft,C.L.,
            Charlab,R., Collins,F.H., Venter,J.C. and Hoffman,S.L.
TITLE       Celeris Anopheles gambiae EST project
JOURNAL     Unpublished (2002)
COMMENT     Contact: Holt R.A.
            Celeris Genomics
            45 W. Gude Dr., Rockville, MD 20850, USA
            Tel: 2404533151
            Fax: 2404534580
            Email: HoltR@celera.com
            Plate: NU010049VR row: N column: 23
            Seq primer: M13 Reverse.

FEATURES
    source
        1..627
        Location/Qualifiers
            /organism="Anopheles gambiae"
            /mol_type="mRNA"
            /strain="RSP-ST (Reduced susc. to Permethrin - std.
            chromosome)"
            /db_xref="taxon:7165"
            /clone="1960049718565"
            /dev_stage="Adult"
            /lab_host="DH10B"
            /clone_1lb="A.Gam.ad.cdna.blood1"
            /note="Vector: pSport1; Site_1: SalI; Site_2: NotI; Whole
            adult mosquitoes (mixed sex) frozen on liquid nitrogen 24
            hours after human blood feeding. cDNA inserts >500 bp
            cloned directionally into pSport 1. Not 1 site is 3'.
            Clones available through the Malaria Research and
            Reference Reagent Resource Center (www.malaria.mri.org)"

ORIGIN
Query Match      71.4%; Score 20; DB 3; Length 627;
Best Local Similarity 82.1%; Pred. No. 3.9e+03;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY      1 GACATTAATGTTAGTTGTATGAGAAC 28
        |||||
        424 GACATTAATGTTAGTTGTATGAGAAC 397

RESULT 31
CZ821000      664 bp      DNA      linear      GSS 26-JUN-2005
LOCUS      CZ821000      664 bp      DNA      linear      GSS 26-JUN-2005
DEFINITION      OC_Ba0199G04.f OC_Ba Oryza coarctata genomic clone OC_Ba0199G04
ACCESSION      CZ821000
VERSION        CZ821000.1 GI:71260853
KEYWORDS      GSS.
SOURCE        Oryza coarctata (Porteresia coarctata)
ORGANISM      Oryza coarctata
REFERENCE     1 (bases 1 to 664)
AUTHORS      Kim,H., Collura,K., Wisotski,M., Byrne,M., Stum,D., Smart,D.,
            Rao,K., Luo,M., Jetly,R., Kudrna,D., Muller,C., Soderlund,C. and
            Wang,R.
TITLE       OMAP (Oryza Map Alignment Project) - Arizona Genomics Institute
JOURNAL     Unpublished (2005)
COMMENT     Contact: Rod A. Wang
            Arizona Genomics Institute
            University of Arizona
            Forbes Building Room 303, Tucson, AZ 85721-0036, USA
            Tel: 520 626 9595
            Fax: 520 621 1259
            Email: rtwing@genome.arizona.edu
            PCR Primers

```

```

FEATURES
    source
        1..664
        Location/Qualifiers
            /organism="Oryza coarctata"
            /mol_type="genomic DNA"
            /db_xref="taxon:77588"
            /clone="OC_Ba0199G04"
            /issue_type="leaves"
            /dev_stage="mature"
            /lab_host="DH10B"
            /clone_1lb="OC_Ba"
            /note="Vector: pGIBBAC1; Site_1: HindIII; Site_2: HindIII"

ORIGIN
Query Match      71.4%; Score 20; DB 13; Length 664;
Best Local Similarity 82.1%; Pred. No. 3.9e+03;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY      1 GACATTAATGTTAGTTGTATGAGAAC 28
        |||||
        492 GACATTAATGTTAGTTGTATGAGAAC 519

RESULT 32
AL849178      665 bp      mRNA      linear      EST 26-NOV-2003
LOCUS      AL849178      665 bp      mRNA      linear      EST 26-NOV-2003
DEFINITION      AL849178 XGC-egg Xenopus tropicalis cDNA clone TEG9002h04 5', mRNA
ACCESSION      AL849178
VERSION        AL849178.2 GI:38560538
KEYWORDS      EST.
SOURCE        Xenopus tropicalis (western clawed frog)
ORGANISM      Xenopus tropicalis
REFERENCE     1 (bases 1 to 665)
AUTHORS      Croining,M.D.R., Ashurst,J.L., Taylor,R., Zorn,A.M. and Rogers,J.
TITLE       Sanger xenopus tropicalis EST project 2001 (11_2003)
JOURNAL     Unpublished (2003)
COMMENT     On Sep 15, 2002 this sequence version replaced gi:22869400.
            Contact: Taylor R
            Sanger Institute
            Hinxton, Cambridgeshire, CB10 1SA, UK
            Email: trop@sanger.ac.uk
            Sanger Xenopus tropicalis EST project 2001
            TROPICALIS SEQUENCE ID: TEG9002h04.PIKSP6
            Sequencing primer: SP6
            This sequence is from a Xenopus Gene Collection (XGC) library
            constructed by Aaron M. Zorn.
            cDNA was oligo dt primed from Sug of poly A+ RNA from egg.
            EcoRI-NotI cut cDNA was then ligated into pCS107 with EcoRI at the
            5' end and NotI at the 3' end.
            Vector: pCS107; Site 1: EcoRI; Site 2: NotI
            Host: Escherichia coli XL1-blue.

FEATURES
    source
        1..665
        Location/Qualifiers
            /organism="Xenopus tropicalis"
            /mol_type="mRNA"
            /db_xref="taxon:8364"
            /clone="TEG9002h04"
            /dev_stage="egg"
            /lab_host="Escherichia coli XL1-blue"
            /note="Vector: pCS107; Site 1: EcoRI; Site 2: NotI; cDNA
            was oligo dt primed from Sug of poly A+ RNA from egg.
            EcoRI-NotI cut cDNA was then ligated into pCS107 with
            EcoRI at the 5' end and NotI at the 3' end"

ORIGIN

```

Query Match 71.4%; Score 20; DB 1; Length 665;
 Best Local Similarity 82.1%; Pred. No. 3.9e+03;
 Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 GACATATTTGTTATGTTTGTATGAAC 28
 |||||
 Db 489 GAAATTATGTTATGTTATGTTATGATAC 516

RESULT 33
 CX689149/c
 LOCUS
 DEFINITION ydc08b02.y2 Sea urchin EST lib1 Strongylocentrotus purpuratus cDNA
 clone ydc08b02 5', mRNA sequence.
 ACCESSION CX689149
 VERSION CX689149
 KEYWORDS EST.
 SOURCE Strongylocentrotus purpuratus
 ORGANISM Strongylocentrotus purpuratus
 Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
 Echinozoa; Euechinozoa; Echinodermata; Echinodermata;
 Strongylocentrotidae; Strongylocentrotus.
 1 (bases 1 to 700)
 Coffman,J.A., Robertson,A.J., Clifton,S., Pape,D., Hillier,L.,
 Martin,J., Wyllie,T., Dante,M., Meyer,R., Theising,B., Bowers,Y.,
 Gibbons,M., Konko,I., Tsagaris,W., Rutter,E., Kennedy,S. and
 Wilson,R.
 Washu Sea Urchin EST Project
 Unpublished (2004)
 Contact: Dr. James A. Coffman
 Washu Sea Urchin EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@wustl.wustl.edu
 DNA sequencing by: Washington University Genome Sequencing Center
 Seq primer: -28RPPOT
 High quality sequence stop: 700.
 Location/Qualifiers
 1..700
 /organism="Strongylocentrotus purpuratus"
 /mol_type="mRNA"
 /db_xref="taxon:7668"
 /clone="ydc08b02"
 /lab_host="DH10B"
 /note="Vector: pCMVSPORT6.1 (Invitrogen); Site 1: NotI;
 Site 2: SmaI; Arrayed normalized library of full-length
 cDNAs representing blastula stage transcriptome of the sea
 urchin Strongylocentrotus purpuratus, cloned into the
 vector pCMVSPORT6.1 (Invitrogen)"

ORIGIN
 Query Match 71.4%; Score 20; DB 9; Length 700;
 Best Local Similarity 82.1%; Pred. No. 3.9e+03;
 Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 GACATATTTGTTATGTTTGTATGAAC 28
 |||||
 Db 533 GACAGTATGCTTACGTTTGTATGAAC 506

RESULT 34
 CX908094
 LOCUS
 DEFINITION JGI CAAN728.fwd NIH XGC tropic4 Xenopus tropicalis cDNA clone
 IMAGE:7686784 5', mRNA sequence.
 ACCESSION CX908094
 VERSION CX908094.1
 KEYWORDS EST.
 SOURCE Xenopus tropicalis (western clawed frog)

ORGANISM
 Xenopus tropicalis
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
 Xenopodidae; Xenopus; Silurana.
 1 (bases 1 to 749)
 Richardson,P., Lucas,S., Rokhsar,D., Dettler,J.C., Ng,D.C.,
 Brokstein,P. and Lindquist,E.A.
 DOE Joint Genome Institute Xenopus tropicalis EST project
 Unpublished (2004)
 Other ESTs: JGI CAAN728.rev
 Contact: Lindquist,E.A., Richardson,P.
 DOE Joint Genome Institute
 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
 Tel: 925 296 5600
 Fax: 925 296 5710
 Email: cna@jgi-psf.org
 Tissue Procurement: Timothy Grammer (Richard M. Harland Laboratory,
 University of California, Berkeley:
 http://tropicalis.berkeley.edu/home)
 cDNA Library Preparation: DOE Joint Genome Institute:
 http://www.jgi.doe.gov
 DNA Sequencing: DOE Joint Genome Institute: http://www.jgi.doe.gov
 Clone Distribution: I.M.A.G.E. Consortium/LNL:
 http://image.lnl.gov
 Naming Conventions: EST name is generated by the concatenation of
 the JGI Clone id and the direction of sequencing. The suffix '.fwd'
 indicates a forward sequencing read of the insert. It does not
 necessarily reflect the orientation of the insert.
 Plate: CAAN 0005 row: P column: 14
 High quality sequence stop: 684.
 Location/Qualifiers
 1..749
 /organism="Xenopus tropicalis"
 /mol_type="mRNA"
 /db_xref="taxon:8364"
 /clone="IMAGE:7686784"
 /sex="male"
 /tissue_type="Testes"
 /dev_stage="Adult"
 /lab_host="ElectroMAX DH10B"
 /clone_lib="NIH XGC tropic4"
 /note="Vector: pCMVSPORT6; Site 1: SalI; Site 2: NotI;
 This library was made from dt primed cDNA and cloned into
 Invitrogen pCMVSPORT6 vector. The work was done at DOE
 Joint Genome Institute. Poly A RNA were primed with 5'
 GACTAGTTCTAGATCGGAG CGGCGCCCTTTTCTTTTCTT 3'. cDNA
 were ligated to SalI adapter (5' TCGACCCAGCGGTCG and
 5'CGAGCGGTCG), digested with NotI, size fractionated in
 1.1% agarose gel electrophoresis and ligated into NotI and
 SalI digested pCMVSPORT6 vector."

ORIGIN
 Query Match 71.4%; Score 20; DB 9; Length 749;
 Best Local Similarity 82.1%; Pred. No. 3.9e+03;
 Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 GACATTTTGTATGTTTGTATGAAC 28
 |||||
 Db 242 GAAATTATGCTTAAATGATGATGAAC 269

RESULT 35
 CZ754866
 LOCUS
 DEFINITION OC_Ba0106B12.f OC_Ba Oryza coarctata genomic clone OC_Ba0106B12
 5', genomic survey sequence.
 ACCESSION CZ754866
 VERSION CZ754866.1
 KEYWORDS GSS.
 SOURCE Oryza coarctata (Porteresia coarctata)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; BCP

REFERENCE 1 (bases 1 to 754)
AUTHORS Kim H., Collura, K., Wisotski, M., Byrne, M., Stum, D., Smart, D., Rao, K., Luo, M., Jetty, R., Kudrna, D., Muller, C., Soderlund, C. and Wing, R.
TITLE OMAF (Oryza Map Alignment Project) - Arizona Genomics Institute
JOURNAL Unpublished (2005)
COMMENT Contact: Rod A. Wing
Arizona Genomics Institute
University of Arizona
Forbes Building Room 303, Tucson, AZ 85721-0036, USA
Tel: 520 626 9555
Fax: 520 621 1259
Email: twing@genome.arizona.edu
PCR Primers
FORWARD: TAA TAC GAC TCA CTA TAG GG
BACKWARD: CAC TCA TTA GGC ACC CCA
Plate: 0106 row: B column: 12
Seq primer: TAA TAC GAC TCA CTA TAG GG
Class: BAC ends.
Location/Qualifiers
1..754
/organism="Oryza coarctata"
/mol_type="genomic DNA"
/db_xref="taxon:77588"
/clone="OC_Ba0106B12"
/issue_type="leaves"
/dev_stage="mature"
/lab_host="DH10B"
/clone_lib="OC_Ba"
/note="Vector: pAG1BAC1; Site_1: HindIII; Site_2: HindIII"

ORIGIN
Query Match 71.4%; Score 20; DB 13; Length 754;
Best Local Similarity 82.1%; Pred. No. 3.9e+03;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 GACATTATGTTAGTTGTATGTAAC 28
Db 597 GATATGAGCTTATAGTTGTTGTATC 624

RESULT 36
DV932465 763 bp mRNA linear EST 01-DEC-2005
LOCUS LB03025_CR_019 GC_BGC-30 Bos taurus cDNA clone IMAGE:8141805, mRNA
DEFINITION sequence.
ACCESSION DV932465
VERSION DV932465.1 GI:82989654
KEYWORDS EST.
SOURCE Bos taurus (cattle)
ORGANISM Bos taurus
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
Pecora; Bovidae; Bovinae; Bos.
1 (bases 1 to 763)
Moore, S., Alexander, L., Brownstein, M., Guan, L., Lobo, S., Meng, Y., Taniguchi, M., Wang, Z., Yu, J., Prange, C., Schreiber, K., Shennen, C., Wagner, L., Bala, M., Barabak, S., Barber, S., Babaloff, R., Beland, J., Chun, E., Del Rio, L., Gibson, S., Hanson, R., Kirpatrick, R., Liu, J., Matsuo, C., Mayo, M., Santos, R., Scott, J., Tsai, M., Wong, D., Siddiqui, A., Holt, R., Jones, S. J. and Marra, M. A.
Bovine Genome Sequencing Program: Full-length cDNA Sequencing
Unpublished (2005)
Contact: Robert Kirkpatrick
Canada's Michael Smith Genome Sciences Centre
BC Cancer Agency
Suite 100, 570 West 7th Avenue, Vancouver, British Columbia,
Canada, V5Z 4S6
Tel: 1-604-707-5900 x5406
Fax: 1-604-876-3561
Email: robertk@bcgsc.ca
Insert Length: 763 Std Error: 0.00

TITLE JOURNAL
COMMENT

Plate: LB03025 row: 0 column: 19
High quality sequence stop: 763.
Location/Qualifiers
1..763
/organism="Bos taurus"
/mol_type="mRNA"
/strain="U1 Hereford"
/db_xref="taxon:9913"
/clone="IMAGE:8141805"
/sex="female"
/issue_type="Muscle"
/dev_stage="6 months old fetus"
/lab_host="E. coli DH10B T1 Phage resistant"
/clone_lib="GC_BGC-30"
/note="Vector: pXpress 1; Site_1: blunt (5' end of cDNA); Site_2: NotI (3' end of cDNA)"

ORIGIN
Query Match 71.4%; Score 20; DB 10; Length 763;
Best Local Similarity 82.1%; Pred. No. 3.9e+03;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 GACATTATGTTAGTTGTATGTAAC 28
Db 703 GCCATTATGTTAGTTGTATGTAAC 730

RESULT 37
CX896534 832 bp mRNA linear EST 04-FEB-2005
LOCUS JGI_CAA6990.fwd NIH XGC_tropes3 Xenopus tropicalis cDNA clone
DEFINITION IMAGE:7680909 5', mRNA sequence.
ACCESSION CX896534
VERSION CX896534.1 GI:58635878
KEYWORDS EST.
SOURCE Xenopus tropicalis (western clawed frog)
ORGANISM Xenopus tropicalis
Eukaryote; Metazoa; Chordata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
Xenopodinae; Xenopus; Silurana.
1 (bases 1 to 832)
Richardson, P., Lucas, S., Rohsard, D., Dettler, J. C., Ng, D. C., Brokstein, P. and Lindquist, E. A.
DOE Joint Genome Institute Xenopus tropicalis EST project
Unpublished (2004)
Other ESTs: JGI_CAA6990.rev
Contact: Lindquist, E. A., Richardson, P.
DOE Joint Genome Institute
2800 Mitchell Drive, Walnut Creek, CA 94598, USA
Tel: 925 296 5600
Fax: 925 296 5710
Email: cdna@jgi-psf.org
Tissue Procurement: Timothy Grammer (Richard M. Harland Laboratory,
University of California, Berkeley)
http://tropicalis.berkeley.edu/home
cDNA Library Preparation: DOE Joint Genome Institute:
http://www.jgi.doe.gov
DNA Sequencing: DOE Joint Genome Institute: http://www.jgi.doe.gov
http://image.llnl.gov
Naming Conventions: EST name is generated by the concatenation of
the JGI Clone id and the direction of sequencing. The suffix '.fwd'
indicates a forward sequencing read of the insert. It does not
necessarily reflect the orientation of the insert.
Plate: CAA6 0073 row: k column: 19
High quality sequence stop: 724.
Location/Qualifiers
1..832
/organism="Xenopus tropicalis"
/mol_type="mRNA"
/db_xref="taxon:8364"
/clone="IMAGE:7680909"
/sex="male"

FEATURES
source

/tissue type="Testes"
 /lab_host="Adult"
 /clone_id="NH_XGC_tropes3"
 /note="Vector: pCMVSPORT6; Site 1: SalI; Site 2: NotI;
 This library was made from dt primed cDNA and cloned into
 Invitrogen pCMVSPORT6 vector. The work was done at DOE
 Joint Genome Institute. Poly A RNA were primed with 5'
 GACTAGTCTTAGATCGGAG CGAGCCGCCCTTTTCTTTTCTTTT 3'. cDNA
 were ligated to SalI adapter (5' TGACCCACGCGTCG and
 5'CGACGCGTCGG), digested with NotI, size fractionated in
 1.1% agarose gel electrophoresis and ligated into NotI and
 SalI digested pCMVSPORT6 vector."

ORIGIN
 Query Match 71.4%; Score 20; DB 9; Length 832;
 Best Local Similarity 82.1%; Pred. No. 3.9e+03;
 Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 GACATTATGTTATGTTGATGGAAC 28
 |||||
 Db 637 GAATTATGCTGTAATAGATGGAAC 664

RESULT 38
 CG920252 954 bp DNA linear GSS 12-DEC-2003
 MBEN064TR mth2 Medicago truncatula genomic clone 10K7, genomic
 survey sequence.
 CG920252
 CG920252.1 GI:39779935
 GSS.
 Medicago truncatula (barrel medic)
 SOURCE
 Medicago truncatula
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifoliaceae;
 Medicago.
 1 (bases 1 to 954)
 Town, C.D., Shetty, J., Koo, H. and Feldblum, T. F.
 Sequencing of BAC ends from Medicago truncatula
 Unpublished (2003)
 Other GSSs: MBEN064TF
 Contact: Chris Town
 TIGR
 9712 Medical Center Drive, Rockville, MD 20850, USA.
 Tel: 301-838-3523
 Fax: 301-838-0208
 Email: cdtown@tigr.org
 Seq primer: CAGGAAACAGCTATGACC
 Class: BAC ends.
 Location/Qualifiers
 1..954
 /organism="Medicago truncatula"
 /mol_type="genomic DNA"
 /cultivar="genotype A17"
 /db_xref="taxon:3880"
 /clone="10K7"
 /clone_1lb="mth2"
 /note="Vector: pBeloBAC11; Site 1, HindIII; Site 2:
 HindIII; Cook, D.R. and Kim, D.J., unpublished"

ORIGIN
 Query Match 71.4%; Score 20; DB 12; Length 954;
 Best Local Similarity 82.1%; Pred. No. 3.9e+03;
 Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 GACATTATGTTATGTTGATGGAAC 28
 |||||
 Db 68 GAAGTATTTTAATAGTTGATGGAAC 95

RESULT 39

CF091813/c
 LOCUS 511 bp mRNA linear EST 22-JUN-2003
 DEFINITION QM99H07.yg.ab1 QM M sunflower H. argophyllus Helianthus argophyllus
 cDNA clone QM99H07, mRNA sequence.
 CF091813
 CF091813 GI:33130880
 EST.
 SOURCE
 Helianthus argophyllus
 Helianthus argophyllus
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 asterids; campanulids; Asterales; Asteraceae; Asteroideae;
 Heliantheae; Helianthus.
 1 (bases 1 to 511)
 Kozik, A., Michelmore, R. W., Knapp, S., Matvienko, M., Rieseberg, L.,
 Lin, H., Van Damme, M., Lavelle, D., Chevalier, P., Ziegler, J.,
 Ellison, P., Kolkmann, J., Slabaugh, M. S., Livingston, K., Zhou, Y.,
 Lai, Z., Church, S., Jackson, L. and Bradford, K.
 Lettuce and Sunflower ESTs from the Compositae Genome Project
 http://compositae.ucdavis.edu/
 Unpublished (2002)
 Contact: Alexander Kozik [R.W.Michelmore]
 Department of Vegetable Crops, R.W.Michelmore Lab
 University of California at Davis (UCD)
 Asmundson Hall, UCD, Davis, CA 95616, USA
 Tel: 1-(530)-742-1742
 Fax: 1-(530)-752-9659
 Email: akozik@ucdavis.edu [michelmoreveg@mail.ucdavis.edu]
 belongs to contig QM_CA_Contig4705, see http://cspdb.ucdavis.edu/
 for details.
 Plate: QM99 row: H column: 07.
 Location/Qualifiers
 1..511
 /organism="Helianthus argophyllus"
 /mol_type="mRNA"
 /db_xref="taxon:73275"
 /clone="QM99H07"
 /lab_host="E.coli"
 /clone_1lb="QM M sunflower H. argophyllus"
 /note="Vector: pBRcNAB51ab; The library was constructed
 from three different sources (seedling, root and leaf) of
 RNA from a single genotype. cDNAs were pooled and
 directionally cloned into a custom medium-copy vector.
 Details of library construction can be obtained at
 http://cspdb.ucdavis.edu/"

ORIGIN
 Query Match 70.7%; Score 19.8; DB 5; Length 511;
 Best Local Similarity 91.3%; Pred. No. 4.6e+03;
 Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 CATTATGTTATGTTGATG 25
 |||||
 Db 306 CAATATGCTATGTTGATGG 284

RESULT 40
 B2392066 536 bp DNA linear GSS 30-APR-2003
 LOCUS E1NCK17P E1 10.12 KB Entamoeba invadens genomic clone E1NCK17,
 genomic survey sequence.
 B2392066
 B2392066.1 GI:30238603
 GSS.
 SOURCE
 Entamoeba invadens
 Entamoeba invadens
 Eukaryota; Entamoebidae; Entamoeba.
 1 (bases 1 to 536)
 Wang, Z., Samuelson, J., Clark, C. G., Eichinger, D., Paul, J., van
 Dellen, K., Hall, N., Anderson, I. and Lotz, B.
 Gene discovery in the Entamoeba invadens genome
 Mol. Biochem. Parasitol. 129 (1), 23-31 (2003)
 12798503

COMMENT Other_GSSs: EINCK17TR
Contact: Brendan Loftus
Department of Eukaryotic Genomics
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-3543
Fax: 301-838-0208
Email: ent@tigr.org
DNA was provided by Daniel Bichinger
Seq primer: TP
Class: sheared ends.

FEATURES
source
1..536
Location/Qualifiers
/organism="Entamoeba invadens"
/mol_type="genomic DNA"
/strain="IP-1"
/db_xref="taxon:33085"
/clone="EINCK17"
/clone_1lb="E1_10_12_KB"
/note="Vector: pHOS2; Site 1: BstXI; Total genomic DNA was isolated from early log phase trophozoites of E. invadens IP-1 using a Qiagen plant DNA extraction kit. A shotgun medium-size plasmid library (average insert size of 10-12 kb) was generated by random mechanical shearing of E. invadens genomic DNA, repairing the ends of DNA fragments with T4 polymerase, adding BstXI adaptors and ligating into the BstXI site of a pUC-derived vector pHOS2."

ORIGIN
Query Match 70.7%; Score 19.8; DB 11; Length 536;
Best Local Similarity 91.3%; Pred. No. 4.6e+03;
Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GACATTATGTTAGTTGTAT 23
|||||
186 GACATTTTATTATTTGTTT 208

Db

RESULT 41
BU026745/c 592 bp mRNA linear EST 23-AUG-2002
LOCUS OHG17014.yg.ab1 QH ERGHJ sunflower RHA280 Helianthus annuus cDNA
DEFINITION OHG17014.yg.ab1 QH ERGHJ sunflower RHA280 Helianthus annuus cDNA
ACCESSION BU026745
VERSION BU026745
KEYWORDS GI:22462265
SOURCE EST.
ORGANISM Helianthus annuus (common sunflower)
Helianthus annuus
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; asterids; campanulids; Asterales; Asteraceae; Asteroideae; Helianthaceae; Helianthus.
1 (bases 1 to 592)
Kozik,A., Michelmore,R.W., Knapp,S., Matvienko,M., Riesberg,L., Lin,H., Van Damme,M., Laveille,D., Chevalier,P., Ziegler,J., Ellison,P., Kolkmann,J., Slabaugh,M.S., Livingston,K., Zhou,Y., Lai,Z., Church,S., Jackson,L. and Bradford,K.
Lecture and Sunflower ESTs from the Compositae Genome Project
http://compogenomics.ucdavis.edu/
unpublished (2002)
Contact: Alexander Kozik (R.W.Michelmore)
Department of Vegetable Crops, R.W.Michelmore Lab
University of California at Davis (UCD)
Aamundson Hall, UCD, Davis, CA 95616, USA
Tel: 1-(530)-742-1742
Fax: 1-(530)-752-9659
Email: akozik@ucdavis.org [michelmore@vegmail.ucdavis.edu]
Singleton, see http://cspdb.ucdavis.edu/ for details.
Plate: OHG17 row: O Column: 14.
Location/Qualifiers
1..592
/organism="Helianthus annuus"
/mol_type="mRNA"

FEATURES
source

COMMENT /cultivar="RHA280"
/db_xref="taxon:4232"
/clone="OHG17014"
/lab_host="E.coli"
/clone_1lb="QH_ERGHJ sunflower RHA280"
/note="Vector: pBRCDNA5flAB; The library was constructed from 11 different sources of RNA from a single genotype. Separate cDNAs were generated using primers that incorporated unique 5' and 3' tags to distinguish each source of RNA. cDNAs were then pooled, size-fractionated, directionally cloned into a custom medium-copy vector and transformations made with four size classes to minimize size bias. Details of each source of RNA and library construction can be obtained at http://cspdb.ucdavis.edu/TAG_LIB-QH_ERGHJ sunflower RHA280
TAG_LIB-QH_ERGHJ sunflower RHA280
TAG_SBO-GCTACTCCGG"

ORIGIN
Query Match 70.7%; Score 19.8; DB 3; Length 592;
Best Local Similarity 91.3%; Pred. No. 4.6e+03;
Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 CATTATGTTATGTTGTATGG 25
|||||
419 CAATATGTCATGTTGTATGG 397

Db

RESULT 42
DX044921/c 770 bp DNA linear GSS 10-JAN-2006
LOCUS DX044921
DEFINITION KBrB046A24R KBrB, Brassica rapa BamHI BAC library Brassica rapa subsp. pekinensis genomic clone KBrB046A24, genomic survey sequence.
ACCESSION DX044921
VERSION DX044921
KEYWORDS GI:84739218
SOURCE GSS.
ORGANISM Brassica rapa subsp. pekinensis
Brassica rapa subsp. pekinensis
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
1 (bases 1 to 770)
Yang,T.J., Kwon,S.J., Kim,J.A., Kim,J.S., Lim,K.B., Jin,M., Park,J.Y., Lim,M.H., Kim,H.I., Choi,B.S., Seol,Y.J., Park,D.S., Hahn,J.H. and Park,B.S.
End sequence of Brassica rapa BamHI (KBrB) BAC clone
Unpublished (2005)
Contact: Beom-Seok Park
Brassica Genomics Team
National Institute of Agricultural Biotechnology
225 Seodun-Dong, Suwon, 441-707, Korea
Tel: +82-31-299-1670
Fax: +82-31-299-1672
Email: pbeom@da.90.kr
BAC end sequence of Brassica rapa ssp. pekinensis BamHI BAC clone KBrB046A24
Seq primer: M13 Reverse
Class: BAC ends.
Location/Qualifiers
1..770
/organism="Brassica rapa subsp. pekinensis"
/mol_type="genomic DNA"
/cultivar="Chifu"
/sub_species="pekinensis"
/db_xref="taxon:51351"
/clone="KBrB046A24"
/lab_host="E.coli DH10B"
/clone_1lb="KBrB, Brassica rapa BamHI BAC library"
/note="Vector: pCUGIBAC1; Site 1: BamHI; Brassica rapa ssp. pekinensis var. Chifu BAC library (KBrB BAC) is provided by Yong-Pyo Lim (CNU)."

FEATURES
source

ORIGIN

Query Match 70.7%; Score 19.8; DB 14; Length 770;
 Best Local Similarity 91.3%; Pred. No. 4.6e+03;
 Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 ACATTATTGTTATGATTGATG 24
 |||||
 248 ACATTATTGTTATGATTGTTG 226

RESULT 43
 CINS07B84/c
 LOCUS
 DEFINITION T7 end of clone BC0A010A01 of library BC0A from strain CBS 767 of
 Debaryomyces hansenii, genomic survey sequence.
 AL437546
 AL437546.1 GI:12220959
 GSS.
 Debaryomyces hansenii var. hansenii
 Debaryomyces hansenii var. hansenii
 Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 Saccharomycetales; Saccharomycetaceae; Debaryomyces.
 1 (bases 1 to 998)
 Souciet,J., Aigle,M., Artiguenave,F., Blandin,G.,
 Bolotin-Fukuhara,M., Bon,E., Brotier,P., Casaregola,S., de
 Montigny,J., Dujon,B., Durrens,P., Galliardin,C., Lepingle,A.,
 Llorente,B., Maupertuy,A., Neuvéglise,C., Ozier-Kalogeropoulos,O.,
 Potier,S., Saurin,W., Tekala,F., Toffano-Nioche,C.,
 Wesolowski-Louvel,M., Wincker,P. and Weissbach,J.
 Genomic exploration of the hemiascomycetous yeasts: 1. A set of
 yeast species for molecular evolution studies
 FEMS Lett. 487 (1), 3-12 (2000)

JOURNAL 1152876
 2 (bases 1 to 998)
 Lepingle,A., Casaregola,S., Neuvéglise,C., Bon,E., Nguyen,H.,
 Artiguenave,F., Wincker,P. and Galliardin,C.
 Genomic exploration of the hemiascomycetous yeasts: 14.
 Debaryomyces hansenii var. hansenii
 FEMS Lett. 487 (1), 82-86 (2000)

REFERENCE 1152889
 3 (bases 1 to 998)
 Genoscope.
 Direct Submission
 Submitted (08-SEP-2000) Genoscope - Centre National de Séquençage,
 2 rue Gaston Crémieux, CP 5706, 91057 EVRY cedex - FRANCE (E-mail :
 seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
 This GSS is part of a random genomic sequencing program of thirteen
 yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces
 exiguus, Saccharomyces servazzii, Zygosaccharomyces rouxii,
 Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces
 lactis var. lactis, Kluyveromyces marxianus var. marxianus, Pichia
 angusta, Debaryomyces hansenii var. hansenii, Pichia sorbitophila,
 Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to
 5 kb were prepared and both extremities were sequenced. See
 keywords for description of this sequence and for the sequence of
 the other extremity of this insert.
 Location/Qualifiers
 1..998
 /organism="Debaryomyces hansenii var. hansenii"
 /mol_type="genomic DNA"
 /strain="CBS 767"
 /variety="hansenii"
 /db_xref="taxon:58641"
 /clone="BC0A010A01"
 /clone_1lb="BC0AA"
 /note="end : T7"

ORIGIN
 Query Match 70.7%; Score 19.8; DB 14; Length 998;
 Best Local Similarity 91.3%; Pred. No. 4.6e+03;
 Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 TTATGTTATGATTGATGAA 27

Db 764 TTATGTTATGATTGATGAA 742
 |||||
 192 ACATTATTGTTATGATTGTTG 170

RESULT 44
 AJ854876/c
 LOCUS
 DEFINITION Brassica rapa subsp. pekinensis GSS, clone KBRH051F18, end read,
 primer M13(forward), genomic survey sequence.
 AJ854876
 AJ854876.1 GI:55651482
 GSS; genome survey sequence.
 KEYWORDS
 SOURCE Brassica rapa subsp. pekinensis
 ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
 1
 Bancroft,I., Huckle,B.J., Trick,M. and Beckett,P.
 BAC end sequence of Brassica rapa
 Unpublished
 2 (bases 1 to 1268)
 Beckett,P.
 Direct Submission
 Submitted (05-NOV-2004) Beckett P., Computational Biology, John
 Innes Centre, Norwich Research Park, Colney, Norwich, Norfolk, NR4
 7UH, UNITED KINGDOM
 Contact: Ian Bancroft
 John Innes Centre Norwich Research Park, Colney, Norwich, NR4 7UH.
 Tel: +44 1603 450843
 Fax: +44 1603 450021
 Email: ian.bancroft@bbsrc.ac.uk
 BAC End sequence of Brassica rapa BAC clone KBRH051F18 Seq primer:
 M13(forward)
 Class: BAC ends.
 Location/Qualifiers
 1..1268
 /organism="Brassica rapa subsp. pekinensis"
 /mol_type="genomic DNA"
 /cultivar="Chilly"
 /sub_species="pekinensis"
 /db_xref="taxon:51351"
 /clone="KBRH051F18"
 /clone_1lb="Brassica rapa BAC library (KBRH), Vector:
 pCG18bcl; Site_1: HindIII"

ORIGIN
 Query Match 70.7%; Score 19.8; DB 14; Length 1268;
 Best Local Similarity 91.3%; Pred. No. 4.7e+03;
 Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 ACATTATTGTTATGATTGATG 24
 |||||
 192 ACATTATTGTTATGATTGTTG 170

RESULT 45
 CK928648/c
 LOCUS
 DEFINITION p5mngcr_004935 Normalized Magnaporthe grisea cDNA p5mngcr_004935, mRNA
 library Magnaporthe grisea cDNA clone p5mngcr_004935, mRNA
 sequence.
 CK928648
 CK928648.1 GI:45416499
 EST.
 Magnaporthe grisea (anamorph: Pyricularia grisea)
 Magnaporthe grisea
 Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 Sordariomycetes incertae sedis; Magnaportheaceae; Magnaporthe.
 1 (bases 1 to 170)
 Chen,B., Li,Y., Peng,Y., Dong,H. and Li,D.
 Large-scale identification of ESTs from Magnaporthe grisea by
 normalized cDNA library sequencing

JOURNAL
COMMENT Unpublished (2004)
 Contact: Baoshan Chen, Youzhi Li
 Laboratory of Subtropical Bioresource Conservation and Utilization
 Guangxi University, China Agricultural University, Zhejiang
 University
 Daxue Road 100#, Nanning, Guangxi, 530004, China
 Tel: 0086-771-3239566
 Fax: 0086-771-3237873
 Email: chenbs@n.gx.cninfo.net, URL: http://www.estarray.org
 Seq primer: M13 forward primer.

FEATURES
 source
 1..170
 /organism="Magnaporthe grisea"
 /mol_type="mRNA"
 /db_xref="taxon:148305"
 /clone="p5imgcr.004935"
 /tissue_type="Mycelium, conidium, germinating conidium,
 swelling appressorium, mature appressorium, penetration
 peg"
 /dev_stage="Mycelium, conidium, germinating conidium,
 swelling appressorium, mature appressorium, penetration
 peg"
 /clone_lib="Normalized Magnaporthe grisea cDNA pGEM-T Easy
 library"
 /note="Vector: pGEM-T Easy"

ORIGIN
 Query Match 70.0%; Score 19.6; DB 5; Length 170;
 Best Local Similarity 84.6%; Pred. No. 5.4e+03;
 Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY
 2 ACATTATTTTATGTTTGTATGGA 27
 52 ACATTGCTGTTATGTTTGTATGGA 27

Db

RESULT 46
 LOCUS BI034170 232 bp mRNA linear EST 14-JUN-2001
 DEFINITION PM2-NN0169-190301-006-b03 NN0169 Homo sapiens cDNA, mRNA sequence.
 ACCESSION BI034170
 VERSION BI034170.1 GI:14440796
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homidae; Homo.
 1 (bases 1 to 232)
 Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
 Nagai,M.A., da Silva,W.Jr., Zago,M.A., Bordin,S., Costa,F.F.,
 Goldman,G.H., Carvalho,A.F., Matsumura,A., Bala,G.S., Simpson,D.H.,
 Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
 O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
 Simpson,A.J.
 Shotgun sequencing of the human transcriptome with ORF expressed
 sequence tags
 Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 10737800
 Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil
 Tel: +55-11-27049922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome
 Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=PM2&ct=PM2-NN0169-
 190301-006-b03&ct3=2001-03-19&ct4=1)
 Seq primer: puc 18 forward
 High quality sequence start: 5

FEATURES
 source High quality sequence stop: 231.
 Location/Qualifiers
 1..232
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /dev_stage="Adult"
 /clone_lib="NN0169"
 /note="Organ: nervous normal; Vector: puc18; Site 1: SmaI;
 site 2: SmaI; A mini-library was made by cloning products
 derived from OREGON PCR (U.S. Letters Patent application
 No. 196,716 - Ludwig Institute for Cancer Research)
 profiles into the pUC 18 vector. Reverse transcription of
 tissue mRNA and cDNA amplification were performed under
 low stringency conditions."

ORIGIN
 Query Match 70.0%; Score 19.6; DB 2; Length 232;
 Best Local Similarity 84.6%; Pred. No. 5.4e+03;
 Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY
 3 CATTATTTTATGTTTGTATGAC 28
 24 CAGTATGTGATCTTGTATGAC 49

Db

RESULT 47
 LOCUS DV017521 297 bp mRNA linear EST 23-SEP-2005
 DEFINITION Zp-e5 LXL Zimnia elegans cDNA clone Zp-e5, mRNA sequence.
 ACCESSION DV017521
 VERSION DV017521.1 GI:76151574
 KEYWORDS EST.
 SOURCE Zimnia elegans
 ORGANISM Zimnia elegans
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 asterids; campanulide; Asterales; Asteraceae; Asteroideae;
 Heliantheae; Zimnia.
 1 (bases 1 to 297)
 Pasquet,E., Ranocha,P., Legay,S., Dignonet,C., Barbier,O.,
 Pichon,M. and Goffner,D.
 Novel markers of xylogenesis in zimnia are differentially regulated
 by auxin and cytokinin
 Plant Physiol. 139 (4), 1821-1839 (2005)
 16306148
 Contact: Deborah Goffner
 Differenciation du xyleme : Signaux cellulaires et dynamique
 parietale
 UMR 5546 CNRS/Universite Paul Sabatier Surfaces cellulaires et
 signalisations chez les vegetaux
 Pile de Biotechnologies Vegetales, 24, chemin de Borde Rouge, B.P.
 17 Ausseville, 31326 CASTANET-TOLOSAN, FRANCE
 Tel: 33 5 62 19 35 23
 Fax: 33 5 62 19 35 02
 Email: goffner@cerv.ups-clse.fr
 EST from differentiating TE SSH library - dehydration-induced
 protein RD22 - singulet
 PCR Primers
 FORWARD: TCGAGCGCGCCGCCGCGGAGGT
 BACKWARD: AGCGTGTCGCGCGGAGGT
 Insert Length: 297 Std Error: 0.00
 Plate: 2p row: e column: 05
 Seq primer: M13-reverse
 POLYA=No.
 Location/Qualifiers
 1..297
 /organism="Zimnia elegans"
 /mol_type="mRNA"
 /cultiyar="Envy"
 /db_xref="taxon:34245"
 /clone="Zp-e5"
 /cell_type="TE differentiating cell cultures"

/dev stage="Late TE differentiation processes (48, 60 and 72h of culture in NAA+BAP medium)"
 /lab_host="B.coli DH5alpha"
 /clone_id="LXL"
 /note="Vector: pGEM-T; Site 1: SalI; Site 2: NdeI; LXL (Late Xylogenesis Library). SSH (suppression subtractive hybridization) library made with tester sample comprising pooled RNA from 48, 60 and 72h of cell culture in differentiation medium (supplemented with auxin NAA and cytokinin BAP) and driver sample comprising pooled RNA from 0h of culture and 48, 60 and 72h in control conditions (with NAA only, with BAP only and with no hormone)"

ORIGIN

Query Match 70.0%; Score 19.6; DB 10; Length 297;
 Best Local Similarity 84.6%; Pred. No. 5.4e+03;
 Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 2 ACATTATTGTTAGTTGTATGGA 27
 |||||
 Db 204 ACATCATGTTATGTTATGGA 179

RESULT 48
 CK553998/c 299 bp mRNA linear EST 05-OCT-2004
 LOCUS rswla0_015679.y1 swl Bombyx mori cDNA, mRNA sequence.
 DEFINITION CK553998
 ACCESSION CK553998.1 GI:40938452
 VERSION
 KEYWORDS EST.

SOURCE Bombyx mori (domestic silkworm)
 ORGANISM Bombyx mori

Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pelegrina;
 Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
 Bombycoidea; Bombycidae; Bombyx.

REFERENCE

1 (bases 1 to 299)

Xia, Q., Zhou, Z., Lu, C., Cheng, D., Dai, F., Li, B., Zhao, P., Zha, X.,
 Cheng, T., Chai, C., Fan, G., Xu, J., Liu, C., Lin, Y., Qian, J., Hou, Y.,
 Wu, Z., Li, G., Pan, M., Li, C., Shen, Y., Lan, X., Yuan, L., Li, T.,
 Xu, H., Yang, G., Wan, Y., Zhu, Y., Yu, M., Shen, W., Wu, D., Xiang, Z.,
 Yu, J., Wang, J., Li, R. Q., Shi, J. P., Li, H., Li, G. Y., Su, J. N.,
 Wang, X. L., Li, G. Q., Zhang, Z. J., Wu, Q. F., Li, J., Zhang, Q. P., Wei, N.,
 Xu, J. Z., Sun, H. B., Dong, L., Liu, D. Y., Zhao, S. L., Zhao, X. L.,
 Meng, Q. S., Lan, F. D., Huang, X. G., Li, Y. Z., Fang, F., Li, C. F.,
 Li, D. W., Sun, Y. Q., Zhang, Z. P., Yang, Z., Huang, Y. Q., Qi, Q. H.,
 He, D. D., Huang, H. Y., Zhang, X. W., Wang, Z. Q., Li, W. J., Cao, Y. Z.,
 Yu, Y., Yu, H., Ji, H., Ye, J., Chen, H., Zhou, Y., Liu, B., Wang, J.,
 Ye, J., Li, H., Li, S., Ni, P., Zhang, J., Zhang, Y., Zheng, H., Mao, B.,
 Mao, B., Ye, C., Li, S., Wang, J., Wang, G. K. and Yang, H.

A draft sequence for the genome of the domesticated silkworm (Bombyx mori)

JOURNAL Science 306 (5703), 1937-1940 (2004)

COMMENT

Contact: Zhonghui Xiang
 Southwest Agricultural University
 Chongqing Beibei
 Tel: 86-23-68251123
 Fax: 86-23-68251128
 Email: xzh@swau.cq.cn.

FEATURES

location/Qualifiers
 1..299
 /organism="Bombyx mori"
 /mol_type="mRNA"
 /strain="Dazhao(P50)"
 /db_xref="taxon:7091"
 /sex="mixed"
 /tissue_type="fat body (pupa)"
 /dev_stage="p0"
 /clone_id="swl"
 /note="Vector: pBluescript II SK(+)"

ORIGIN

Query Match 70.0%; Score 19.6; DB 5; Length 299;
 Best Local Similarity 84.6%; Pred. No. 5.4e+03;
 Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 2 ACATTATTGTTAGTTGTATGGA 27
 |||||
 Db 100 ACATTATTGTTAGTTGTATGGA 75

RESULT 49
 BJT24011 305 bp mRNA linear EST 09-MAR-2004
 LOCUS BJT24011 MF01FPA cDNA Oryzias latipes cDNA clone MF01FPA031d06 3',
 DEFINITION mRNA sequence.

ACCESSION BJT24011.1 GI:45265103
 VERSION
 KEYWORDS EST.
 SOURCE Oryzias latipes (Japanese medaka)
 ORGANISM Oryzias latipes

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 Acanthomorpha; Acanthopterygii; Perciformes; Atherinomorpha;
 Belontiiformes; Adrianichthyidae; Oryziinae; Oryzias.

REFERENCE 1 (bases 1 to 305)
 Kohara, Y., Shin-i, T., Kimura, T., Narita, T., Uindo, T. and Takeda, H.
 Medaka EST Project in Takeda's lab
 Unpublished (2001)
 CONTACT: Tadasu Shin-i
 CENTER FOR GENETIC RESOURCE INFORMATION
 NATIONAL INSTITUTE OF GENETICS
 111 Yata, Mishima, Shizuoka 411-8540, Japan
 TEL: 81-559-81-6856
 FAX: 81-559-81-6855
 EMAIL: tshini@gene.nig.ac.jp.

FEATURES

source

/organism="Oryzias latipes"
 /mol_type="mRNA"
 /strain="Md-R"
 /db_xref="taxon:8090"
 /clone="MF01FPA031d06"
 /sex="mixture of female and male"
 /tissue_type="whole embryo"
 /dev_stage="fry stage 40"
 /clone_id="MF01FPA cDNA"

ORIGIN

Query Match 70.0%; Score 19.6; DB 2; Length 305;
 Best Local Similarity 81.5%; Pred. No. 5.4e+03;
 Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 2 ACATTATTGTTAGTTGTATGGAAC 28
 |||||
 Db 6 ACATTCTGCATGTTGTTGATGGAAC 32

RESULT 50
 BM857946 340 bp mRNA linear EST 07-MAR-2002
 LOCUS f417a03.y1 zebrafish sjd day 3 embryo Danio rerio cDNA clone
 DEFINITION IMAGE:5620108 5', mRNA sequence.

ACCESSION BM857946.1 GI:19225628
 VERSION
 KEYWORDS EST.
 SOURCE Danio rerio (zebrafish)
 ORGANISM Danio rerio

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
 Cypriniformes; Cyprinidae; Danio.
 REFERENCE 1 (bases 1 to 340)
 Clark, M., Johnson, S. L., Lehrach, H., Lee, R., Li, F., Marra, M.,
 Eddy, S., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T.,
 Underwood, K., Stepien, M., Theising, B., Allen, M., Bowers, Y.,

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 24, 2006, 05:34:58 ; Search time 342.899 Seconds

(without alignments)
650.664 Million cell updates/sec

Title: us-10-601-913-121

Perfect score: 32

Sequence: 1 TTATTAATAGTGCGCTGCGGTGCCAGAAACC 32

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 5244920 seqs, 3486124231 residues

Total number of hits satisfying chosen parameters: 10489840

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

N_Geneseq.8:.*
1: geneseqn1980s:.*
2: geneseqn1980s:.*
3: geneseqn2000s:.*
4: geneseqn2001as:.*
5: geneseqn2001bs:.*
6: geneseqn2002as:.*
7: geneseqn2002bs:.*
8: geneseqn2003as:.*
9: geneseqn2003bs:.*
10: geneseqn2003cs:.*
11: geneseqn2003ds:.*
12: geneseqn2004as:.*
13: geneseqn2004bs:.*
14: geneseqn2005s:.*
15: geneseqn2006s:.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	32	100.0	32	2	AAT72389 Human Pap
2	32	100.0	32	2	AAT72397 Human Pap
3	32	100.0	32	2	AAT72398 Human Pap
C 4	32	100.0	32	2	AAT72390 Human Pap
5	32	100.0	476	2	AAQ97848 HPV-18 B6
6	32	100.0	476	2	AAZ27568 HPV-18 B6
7	32	100.0	476	2	AAQ61623 CDNA enco
8	32	100.0	477	10	AAU54443 HPV18 B6
9	32	100.0	477	12	AD044101 Nucleotid
10	32	100.0	477	14	AEA51111 Nucleotid
11	32	100.0	483	2	AAQ75471 HPV18 B6/
12	32	100.0	802	1	AAN91786 DNA probe
13	32	100.0	817	2	AAQ29390 DNA enco
14	32	100.0	837	2	AAK78800 HPV fusio
15	32	100.0	837	2	AAK29789 Prot. DI/3
16	32	100.0	837	14	AED52652 Fusion pr
17	32	100.0	843	14	ADV85648 Human pap
18	32	100.0	1000	2	AAQ08623 HPV-18 fr

19	32	100.0	1152	2	AAK78801 HPV fusio
20	32	100.0	1152	2	AAK29790 Prot. DI/3
21	32	100.0	1152	14	AED52656 Fusion pr
22	32	100.0	1731	1	AAN80161 Human pap
23	32	100.0	7857	12	ADZ45648 Human pap
24	30.4	95.0	477	14	AD044106 Nucleotid
25	24	75.0	477	12	AD044105 Nucleotid
26	24	75.0	477	12	AD044112 Nucleotid
27	24	75.0	1095	2	AAZ31781 HPV68 DNA
28	24	75.0	1108	2	AAZ31783 HPV68 B6
29	24	75.0	3283	2	AAZ31789 HPV70 DNA
30	24	75.0	7833	2	AAQ25937 Human pap
31	24	75.0	7905	14	ADZ45655 Human pap
32	23.8	74.4	630	6	ABK62613 Rat seque
33	23.8	74.4	630	13	ADV39914 Rat cardi
34	23.8	74.4	885	4	AAK31153 Human dia
35	23.8	74.4	40	2	AAT71196 HPV type
36	23.2	72.5	30	2	AAQ44290 Sequence
37	23	71.9	30	2	AAQ44290 Sequence
38	23	71.9	30	2	AAZ07822 Oligo 2PV
39	23	71.9	30	14	ADZ64772 HPV 18 B6
40	22.6	70.6	167	5	ABV07997 Human pro
41	22.2	69.4	159	12	ACH89313 Human gen
42	22.2	69.4	297	5	AAK78228 DNA enco
43	22.2	69.4	421	8	ABX40477 Bovine ES
44	22.2	69.4	460	4	AAI10129 Probe #62
45	22.2	69.4	460	4	ABK51759 Human foe
46	22.2	69.4	460	4	AAI31375 Probe #61
47	22.2	69.4	460	4	ABA21587 Probe #53
48	22.2	69.4	460	4	AAK25505 Human bon
49	22.2	69.4	460	4	AAK00066 Human bra
50	22.2	69.4	460	4	ABK25073 Human liv
51	22.2	69.4	460	5	AAI00070 Probe #61
52	22.2	69.4	460	6	ABK00071 Human gen
53	22.2	69.4	581	14	AEI31480 Hamster S
54	22.2	69.4	593	12	ACH75587 Human gen
55	22.2	69.4	659	4	AAK31155 Human dia
56	22.2	69.4	750	4	AAI19425 Probe #93
57	22.2	69.4	750	4	ABK64438 Human foe
58	22.2	69.4	750	4	AAI44613 Probe #13
59	22.2	69.4	750	4	ABA46575 Human bre
60	22.2	69.4	750	4	ABA31574 Probe #10
61	22.2	69.4	750	4	AAK38624 Human bon
62	22.2	69.4	750	4	AAK12895 Human bra
63	22.2	69.4	750	4	ABK38189 Human liv
64	22.2	69.4	750	5	AAI05150 Probe #51
65	22.2	69.4	828	6	ABK12691 Human gen
66	22.2	69.4	925	6	ABN74633 Bovine em
67	22.2	69.4	925	6	ABN74632 Bovine em
68	22.2	69.4	925	6	ABV25873 Human pro
69	22.2	69.4	1023	5	ABV25550 Human pro
70	22.2	69.4	1023	5	ABV25127 Human pro
71	22.2	69.4	1023	5	ABV25127 Human pro
72	22.2	69.4	1023	5	ABV25127 Human pro
73	22.2	69.4	1023	5	ABV25127 Human pro
74	22.2	69.4	1023	5	ABV25127 Human pro
75	22.2	69.4	1023	5	ABV25127 Human pro
76	22.2	69.4	1023	5	ABV25127 Human pro
77	22.2	69.4	1023	5	ABV25127 Human pro
78	22.2	69.4	1023	5	ABV25127 Human pro
79	22.2	69.4	1023	5	ABV25127 Human pro
80	22.2	69.4	1023	5	ABV25127 Human pro
81	22.2	69.4	1023	5	ABV25127 Human pro
82	22.2	69.4	1023	5	ABV25127 Human pro
83	22.2	69.4	1023	5	ABV25127 Human pro
84	22.2	69.4	1023	5	ABV25127 Human pro
85	22.2	69.4	1023	5	ABV25127 Human pro
86	22.2	69.4	1023	5	ABV25127 Human pro
87	22.2	69.4	1023	5	ABV25127 Human pro
88	22.2	69.4	1023	5	ABV25127 Human pro
89	22.2	69.4	1023	5	ABV25127 Human pro
90	22.2	69.4	1023	5	ABV25127 Human pro
91	22.2	69.4	1023	5	ABV25127 Human pro

822	17.2	53.8	38564	13	ADT77132	Adt77132 Type II d	c 895	17	53.1	867	8	ACA48218	ACA48218 Prokaryot
c 823	17.2	53.8	43009	11	ACNA4902	Acna4902 Human gen	c 896	17	53.1	999	10	ABE65433	ABE65433 Rice geno
c 824	17.2	53.8	47756	12	ADOS9512	Ados9512 Human can	c 897	17	53.1	1140	14	ADDA3913	ADDA3913 Chlamydia
c 825	17.2	53.8	57036	14	ADZ13891	Adz13891 Human can	c 898	17	53.1	1140	14	ABE19205	ABE19205 Chlamydia
826	17.2	53.8	57144	4	AAK79963	AAk79963 Human imm	899	17	53.1	1187	13	ADX12729	ADx12729 Plant ful
827	17.2	53.8	62169	12	ADOS9521	Ados9521 Human can	c 900	17	53.1	1205	7	ADBS30984	ADb330984 Human gen
828	17.2	53.8	63529	14	ADOS9521	Ados9521 Human can	c 901	17	53.1	1205	7	ADY36372	ADy36372 HIRA geno
829	17.2	53.8	71108	11	ACNA4782	Acna4782 Murine ca	c 902	17	53.1	1335	13	ADT16701	ADt16701 Plant cdn
830	17.2	53.8	91617	14	ADZ13161	Adz13161 Human can	c 903	17	53.1	1455	13	AAOS167	AAOS167 Taxus cus
831	17.2	53.8	110000	4	AAI99682_07	Contiuation (8 of	c 904	17	53.1	1455	14	ADM72667	Adm72667 Taxus cus
832	17.2	53.8	110000	4	AAI99683_07	Contiuation (8 of	c 905	17	53.1	1494	13	ADK48892	Adk48892 Plant ful
c 833	17.2	53.8	110000	6	ABA90521_06	Contiuation (7 of	c 906	17	53.1	1620	15	ADR31377	Adr31377 Scress re
c 834	17.2	53.8	110000	10	ADBE1169_1	Contiuation (2 of	c 907	17	53.1	1620	15	AEF67886	Aef67886 Methyloph
835	17.2	53.8	110000	10	ADH69807_0	Contiuation (3 of	c 908	17	53.1	2319	13	ADX09399	Adx09399 Plant ful
836	17.2	53.8	110000	10	ACRF7367_19	Contiuation (20 o	c 909	17	53.1	2678	4	ABL11275	ABl11275
837	17.2	53.8	110000	10	ACRF7367_38	Contiuation (39 o	c 910	17	53.1	3009	4	ABL26771	ABl26771 Drosophil
c 838	17.2	53.8	110000	10	ACRF65388_09	Contiuation (10 o	c 911	17	53.1	3325	4	AAK86073	AAk86073 Human imm
839	17.2	53.8	110000	10	ABE279565_2	Contiuation (3 of	c 912	17	53.1	3974	10	ADBS7947	ADb57947 Toxicity-
840	17.2	53.8	110000	12	ADH69807_0	Adh69807 Human Slt	c 913	17	53.1	3974	10	ADBS2437	ADb52437 Toxicity x
841	17.2	53.8	110000	12	ADH69807_0	Adh69807 Human Slt	c 914	17	53.1	3974	10	ABT41756	ABt41756 Toxicity
842	17.2	53.8	110000	14	ADZ13747_1	Contiuation (2 of	c 915	17	53.1	3974	11	ADM21835	Adm21835 Rat hepat
c 843	17.2	53.8	110000	14	ABE42401_31	Contiuation (32 o	c 916	17	53.1	3974	12	ADP72618	Adp72618 Renal tox
c 844	17.2	53.8	110000	14	ABE42737_17	Contiuation (18 o	c 917	17	53.1	3974	13	ADV40836	Adv40836 Rat card
845	17.2	53.8	110000	14	AED76206_08	Contiuation (9 of	c 918	17	53.1	4383	13	ADRO6843	Adro6843 Full leng
846	17.2	53.8	110000	14	AED76206_09	Contiuation (10 o	c 919	17	53.1	5293	4	ABL26770	ABl26770 Drosophil
c 847	17.2	53.8	117143	12	ADJ74882	Adj74882 Marker ge	c 920	17	53.1	8003	8	AAK95877	AAk95877 Human imm
c 848	17.2	53.8	117143	14	AED18148	Aed18148 Fldrotic	c 921	17	53.1	8243	8	ABZ74553	ABz74553 Secreted
c 849	17.2	53.8	117143	15	AEF75107	Aef75107 Human pol	c 922	17	53.1	8243	8	ADPA4545	ADa44545 Human sec
c 850	17.2	53.8	123957	14	ABE839166	Aeb39166 L. pneumo	c 923	17	53.1	10531	4	ABL11274	ABl11274 Drosophil
851	17.2	53.8	125466	13	ABD32897	Abd32897 Human can	c 924	17	53.1	12710	4	AAK81012	AAk81012 Human imm
c 852	17.2	53.8	130001	12	ADJ96282	Adj96282 Human bre	c 925	17	53.1	12710	8	ABZ74552	ABz74552 Secreted
c 853	17.2	53.8	130001	14	ADN98070	Adn98070 Human bre	c 926	17	53.1	12710	8	ADPA4544	ADa44544 Human sec
c 854	17.2	53.8	134292	8	ACA64895	Ac64895 Human GAB	c 927	17	53.1	13842	3	AAZ87297	Aaz87297 S. venezu
c 855	17.2	53.8	134292	11	AED18093	Aed18093 Fldrotic	c 928	17	53.1	13842	12	ADL91915	ADl91915 Streptom
c 856	17.2	53.8	149034	11	AEDP5433	Aedp5433 Human ABC	c 929	17	53.1	15042	4	ABL22246	ABl22246 Drosophil
c 857	17.2	53.8	150481	14	AED89397	Aed89397 Human bre	c 930	17	53.1	16467	10	ACF67365	ACf67365 Phototrab
c 858	17.2	53.8	152501	12	ADP67269	Adp67269 Human chr	c 931	17	53.1	19732	4	AAK89788	AAk89788 Human dig
c 859	17.2	53.8	159321	14	AEB02408	Aee02408 Human her	c 932	17	53.1	22268	6	ABO88219	ABO88219 Human ost
c 860	17.2	53.8	159321	14	AEB02470	Aee02470 Human her	c 933	17	53.1	22909	14	ADZ13860	ADz13860 Murine ca
c 861	17.2	53.8	161573	14	ADY51161	Ady51161 HHV-6 DNA	c 934	17	53.1	24601	2	AAK13160	AAk13160 Enterococ
862	17.2	53.8	161573	14	AEB02552	Aee02552 Human her	c 935	17	53.1	24601	6	ABE89855	ABe89855 Enterococ
c 863	17.2	53.8	161573	14	AEB02604	Aee02604 Human her	c 936	17	53.1	25988	4	ABL09406	ABl09406 Drosophil
864	17.2	53.8	170489	13	ADZ13314	Adz13314 Human can	c 937	17	53.1	27125	12	ADOS9491	ADos9491 Human can
865	17.2	53.8	170506	14	ADZ133520	Adz133520 Human can	c 938	17	53.1	36778	3	AAZ87318	Aaz87318 S. venezu
866	17.2	53.8	171162	14	AED89398	Aed89398 Human bre	c 939	17	53.1	36778	12	ADL91923	ADl91923 Streptom
c 867	17.2	53.8	175077	11	ACNA4626	Acna4626 Human gen	c 940	17	53.1	37948	3	AAZ87285	Aaz87285 S. venezu
c 868	17.2	53.8	183999	13	ADU76470	Adu76470 Human ABC	c 941	17	53.1	38506	3	AAZ56001	AAz56001 Recombina
c 869	17.2	53.8	188267	14	ADZ13837	Adz13837 Human can	c 942	17	53.1	38506	8	ADA09418	ADa09418 Cosmid pk
870	17.2	53.8	188267	12	ADOS9476	Ados9476 Human can	c 943	17	53.1	38506	10	ADH53462	ADH53462 S. venezu
c 871	17.2	53.8	194186	12	ADQ97418	Adq97418 Mouse can	c 944	17	53.1	38506	10	ABE85690	ABe85690 S. venezu
c 872	17.2	53.8	225734	12	ADQ959377	Adq959377 Human can	c 945	17	53.1	40611	12	ADQ97540	Adq97540 Mouse can
c 873	17.2	53.8	225734	14	ADZ13617	Adz13617 Murine ca	c 946	17	53.1	49965	11	ACNA4460	ACna4460 Mouse gen
c 874	17.2	53.8	225734	14	ADZ13617	Adz13617 Murine ca	c 947	17	53.1	49965	11	ACNA4460	ACna4460 Mouse gen
875	17.2	53.8	237326	2	AAV57903	AAv57903 Hereditar	c 948	17	53.1	70931	14	ADZ12976	ADz12976 Murine ca
c 876	17.2	53.8	260209	6	ABE56564	ABe56564 Human SUTL	c 949	17	53.1	73195	10	ACFE5375	ACf56375 Phototrab
c 877	17.2	53.8	260209	12	ADN16204	Adn16204 Human sul	c 950	17	53.1	79084	12	ADQ97563	ADq97563 Human can
c 878	17.2	53.8	267156	6	ABL68560	ABl68560 Kidney ca	c 951	17	53.1	80321	11	ACNA5012	ACna5012 Mouse gen
c 879	17.2	53.8	276276	11	ACNA4350	Acna4350 Human gen	c 952	17	53.1	89014	14	ABE77360	ABe77360 Human TGF
c 880	17.2	53.8	312071	14	AEB04926	Aeb04926 Cancer-as	c 953	17	53.1	89014	14	ABE77360	ABe77360 Human TGF
881	17.2	53.8	347814	12	ADOS9440	Ados9440 Human can	c 954	17	53.1	110000	2	AAZ01425_06	AAZ01425_06
882	17.2	53.1	347814	33	AAK89188	AAk89188 Seg ID No	c 955	17	53.1	110000	6	ABX08336_12	ABX08336_12
883	17.2	53.1	33	2	AAK14994	AAk14994 Oligonuc	c 956	17	53.1	110000	6	ABX08336_12	ABX08336_12
884	17.2	53.1	115	6	AAZ21860	AAz21860 SEAP btop	c 957	17	53.1	110000	12	ADU75985_12	ADU75985_12
885	17.2	53.1	115	6	AAZ21860	AAz21860 SEAP btop	c 958	17	53.1	110000	12	ADU75985_12	ADU75985_12
886	17.2	53.1	115	10	AAZ21860	AAz21860 SEAP btop	c 959	17	53.1	110000	12	ADU75985_12	ADU75985_12
c 887	17.2	53.1	300	6	ABZ14753	ABz14753 Human gen	c 960	17	53.1	110000	12	ADU75985_12	ADU75985_12
888	17.2	53.1	423	6	ABZ14753	ABz14753 Human gen	c 961	17	53.1	110000	13	ABD32806_2	ABD32806_2
889	17.2	53.1	423	6	ABZ14753	ABz14753 Human gen	c 962	17	53.1	110000	14	ABE839175_18	ABE839175_18
890	17.2	53.1	501	6	ABZ14753	ABz14753 Human gen	c 963	17	53.1	110000	14	ABE839175_18	ABE839175_18
891	17.2	53.1	501	6	ABZ14753	ABz14753 Human gen	c 964	17	53.1	110000	14	ABE839175_18	ABE839175_18
c 892	17.2	53.1	597	10	ACF67386	ACf67386 Phototrab	c 965	17	53.1	110000	14	ABE839175_18	ABE839175_18
c 893	17.2	53.1	732	8	ACA33098	ACa33098 Prokaryot	c 966	17	53.1	110000	14	ABE839175_18	ABE839175_18
c 894	17.2	53.1	845	4	AAK63853	AAk63853 Human imm	c 967	17	53.1	122888	6	ABK83569	ABK83569

```

C 968 17 53.1 170469 13 ABD33314 Abd33314 Human can
C 969 17 53.1 170506 14 ADZ13520 Adz13520 Human can
C 970 17 53.1 176594 13 ABD33387 ABD33387 Murine ca
C 971 17 53.1 180000 15 AER49487 AER49487 Marfens'
C 972 17 53.1 180000 15 AER49487 AER49487 Marfens'
C 973 16.8 52.5 66 Adw14220 Adw14220 Human bas
C 974 16.8 52.5 18 3 AAA69132 AAA69132 Bacterioph
C 975 16.8 52.5 243 7 AD667083 AD667083 Corn seed
C 976 16.8 52.5 327 5 AAh68668 AAh68668 Human ant
C 977 16.8 52.5 327 8 Aca99325 Aca99325 Recombina
C 978 16.8 52.5 327 9 Acd45332 Acd45332 Anti-Rh(D
C 979 16.8 52.5 327 15 Aee74495 Aee74495 Anti-Rh(D
C 980 16.8 52.5 330 4 Aa603518 Aa603518 DNA encod
C 981 16.8 52.5 333 4 Aah42399 Aah42399 Nucleotid
C 982 16.8 52.5 466 9 Ach23299 Ach23299 Human adu
C 983 16.8 52.5 470 5 Abv56062 Abv56062 Human pro
C 984 16.8 52.5 565 6 Abk62800 Abk62800 Rat seque
C 985 16.8 52.5 614 13 Acm45597 Acm45597 Cotton pr
C 986 16.8 52.5 624 12 Adn74362 Adn74362 Thale cre
C 987 16.8 52.5 665 13 Adt18740 Adt18740 Plant cDN
C 988 16.8 52.5 673 7 Adr41251 Adr41251 Human CD-
C 989 16.8 52.5 704 3 Aaz65342 Aaz65342 Human sec
C 990 16.8 52.5 704 8 Acc50506 Acc50506 Human sec
C 991 16.8 52.5 704 8 Abz71280 Abz71280 Secreted
C 992 16.8 52.5 704 9 Adb91186 Adb91186 Human sec
C 993 16.8 52.5 704 10 Adc73566 Adc73566 Human sec
C 994 16.8 52.5 704 10 Adr11731 Adr11731 Human sec
C 995 16.8 52.5 708 10 Adc90401 Adc90401 E. faeciu
C 996 16.8 52.5 726 14 Aed12469 Aed12469 Anti-Nogo
C 997 16.8 52.5 741 10 Adg30518 Adg30518 Human GMB
C 998 16.8 52.5 744 10 Aed12432 Aed12432 Anti-Nogo
C 999 16.8 52.5 747 6 Ahs77115 Ahs77115 Frog embr
C1000 16.8 52.5 747 14 Aed12478 Aed12478 Anti-Nogo

```

ALIGNMENTS

```

RESULT 1
AAT72389/c
ID AAT72389 standard; DNA; 32 BP.
XX
AC AAT72389;
XX
DT 09-FEB-1998 (first entry)
XX
DE Human Papillomavirus Type 18 target region.
XX
KW Human Papillomavirus; probe; target region; genital cancer; HPV;
XX cervical smear; ss.
XX
OS Human papillomavirus.
XX
PN BP774518-A2.
XX
PD 21-MAY-1997.
XX
PF 15-NOV-1996; 96EP-00308264.
XX
PR 15-NOV-1995; 95US-0006854P.
XX
PA (GENP-) GEN-PROBE INC.
XX
PI Gordon F, Brentano ST, Carter NM, Hammond PW;
XX WPI; 1997-274349/25.
XX
PT Probes for detection of Human Papillomavirus Type 16 and Type 18 - can
XX distinguish between Type 16 and 18, associated with genital cancers.
XX
PS Claim 16; Page 61; 70pp; English.
XX
CC Novel hybridisation assay probes have been developed comprising an

```

```

CC oligonucleotide which will hybridise under selected conditions to Human
CC Papillomavirus (HPV) Type 16 and/or 18 (but not Types 6, 11, 31, 33, 35,
CC 39, 45, 51, 52, or 58) target nucleic acids to form detectable target;
CC probe duplex. The present sequence represents a specifically claimed
CC target sequence. Oligonucleotides are useful to detect HPV Type 16 and/
CC or 18 in samples e.g. cervical smears, body fluid, and distinguish these
CC from other HPV variants. Papillomaviruses are small DNA viruses and HPV
CC Types 16 and 18 are associated with genital cancers. HPV PCR primers can
CC amplify HPV Type 16 and/or 18 nucleic acid in a sample. HPV Type 16
CC and/or 18 can be detected by adding a probe and detecting probe:target
CC duplex formation; target nucleic acid is optionally amplified. Type 16 or
CC 18 can be specifically detected by amplifying nucleic acids with at least
CC one specifically claimed PCR primer. For Type 18 detection, a helper
CC probe may be used. Oligonucleotides can distinguish target sequences from
CC phylogenetically close HPV Types 16 and 18 whilst available detection
CC sera react with antigens of all papillomaviruses
XX
SQ Sequence 32 BP; 8 A; 8 C; 7 G; 9 T; 0 U; 0 Other;
Query Match 100.0%; Score 32; DB 2; Length 32;
Best Local Similarity 100.0%; Pred. No. 9.2e-05;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TTATTATTAAGGTGCTGCGGTCCGAGAAC 32
DB 32 TTATTATTAAGGTGCTGCGGTCCGAGAAC 1
RESULT 2
AAT72397
ID AAT72397 standard; DNA; 32 BP.
XX
AC AAT72397;
XX
DT 09-FEB-1998 (first entry)
XX
DE Human Papillomavirus Type 18 helper probe.
XX
KW Human Papillomavirus; probe; target region; genital cancer; HPV;
XX cervical smear; ss.
XX
OS Synthetic.
XX
PN BP774518-A2.
XX
PD 21-MAY-1997.
XX
PF 15-NOV-1996; 96EP-00308264.
XX
PR 15-NOV-1995; 95US-0006854P.
XX
PA (GENP-) GEN-PROBE INC.
XX
PI Gordon F, Brentano ST, Carter NM, Hammond PW;
XX WPI; 1997-274349/25.
XX
PT Probes for detection of Human Papillomavirus Type 16 and Type 18 - can
XX distinguish between Type 16 and 18, associated with genital cancers.
XX
PS Claim 20; Page 60; 70pp; English.
XX
CC Novel hybridisation assay probes have been developed comprising an
CC oligonucleotide which will hybridise under selected conditions to Human
CC Papillomavirus (HPV) Type 16 and/or 18 (but not Types 6, 11, 31, 33, 35,
CC 39, 45, 51, 52, or 58) target nucleic acids to form detectable target;
CC probe duplex. The present sequence represents a specifically claimed
CC helper probe. Oligonucleotides are useful to detect HPV Type 16 and/
CC or 18 in samples e.g. cervical smears, body fluid, and distinguish these
CC from other HPV variants. Papillomaviruses are small DNA viruses and HPV
CC Types 16 and 18 are associated with genital cancers. HPV PCR primers can
CC amplify HPV Type 16 and/or 18 nucleic acid in a sample. HPV Type 16

```

CC and/or 18 can be detected by adding a probe and detecting probe:target
CC duplex formation; target nucleic acid is optionally amplified. Type 16 or
CC 18 can be specifically detected by amplifying nucleic acids with at least
CC one specifically claimed PCR primer. For Type 18 detection, a helper
CC probe may be used. Oligonucleotides can distinguish target sequences from
CC phylogenetically close HPV Types 16 and 18 whilst available detection
CC sera react with antigens of all papillomaviruses
XX
SQ Sequence 32 BP; 9 A; 7 C; 8 G; 8 T; 0 U; 0 Other;

Query Match 100.0%; Score 32; DB 2; Length 32;
Best Local Similarity 100.0%; Pred. No. 9.2e-05;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DQ 1 TTATTAATAGTGCTGCGTGCCAGAAACC 32
1 TTATTAATAGTGCTGCGTGCCAGAAACC 32

RESULT 3
AAT72398
ID AAT72398 standard; RNA; 32 BP.
XX AAT72398;
XX
DT 09-FEB-1998 (first entry)
XX
DE Human Papillomavirus Type 18 helper probe.
XX Human Papillomavirus; probe; target region; genital cancer; HPV;
KM cervical smear; ss.
XX
OS Synthetic.
OS Human papillomavirus.
XX
PN EP74518-A2.
XX
PD 21-MAY-1997.
XX
PF 15-NOV-1996; 96BP-00308264.
XX
PR 15-NOV-1995; 95US-0006854P.
XX
PS (GENP-) GEN-PROBE INC.
PA Gordon P, Brentano ST, Carter NM, Hammond PW;
PI
DR WPI; 1997-274349/25.
XX
PT Probes for detection of Human Papillomavirus Type 16 and Type 18 - can
PT distinguish between Type 16 and 18, associated with genital cancers.
XX
PS Claim 20; Page 61; 70pp; English.
XX
CC Novel hybridisation assay probes have been developed comprising an
CC oligonucleotide which will hybridise under selected conditions to Human
CC Papillomavirus (HPV) Type 16 and/or 18 (but not Types 6, 11, 31, 33, 35,
CC 39, 45, 51, 52, or 58) target nucleic acids to form detectable target;
CC probe duplex. The present sequence represents a specifically claimed
CC probe. Oligonucleotides are useful to detect HPV Type 16 and/or
CC 18 in samples e.g. cervical smears, body fluid, and distinguish these
CC from other HPV variants. Papillomaviruses are small DNA viruses and HPV
CC Types 16 and 18 are associated with genital cancers. HPV PCR primers can
CC amplify HPV Type 16 and/or 18 nucleic acid in a sample. HPV Type 16
CC and/or 18 can be detected by adding a probe and detecting probe:target
CC duplex formation; target nucleic acid is optionally amplified. Type 16 or
CC 18 can be specifically detected by amplifying nucleic acids with at least
CC one specifically claimed PCR primer. For Type 18 detection, a helper
CC probe may be used. Oligonucleotides can distinguish target sequences from
CC phylogenetically close HPV Types 16 and 18 whilst available detection
CC sera react with antigens of all papillomaviruses
XX
SQ Sequence 32 BP; 9 A; 7 C; 8 G; 0 T; 8 U; 0 Other;

Query Match 100.0%; Score 32; DB 2; Length 32;
Best Local Similarity 75.0%; Pred. No. 9.2e-05;
Matches 24; Conservative 8; Mismatches 0; Indels 0; Gaps 0;

DQ 1 TTATTAATAGTGCTGCGTGCCAGAAACC 32
1 UUAUUAUAAAGUGCGCGGUGCCAGAAACC 32

RESULT 4
AAT72390/C
ID AAT72390 standard; RNA; 32 BP.
XX
AC AAT72390;
XX
DT 09-FEB-1998 (first entry)
XX
DE Human Papillomavirus Type 18 target region.
XX
KM Human Papillomavirus; probe; target region; genital cancer; HPV;
KM cervical smear; ss.
XX
OS Human papillomavirus.
OS
PN EP74518-A2.
XX
PD 21-MAY-1997.
XX
PF 15-NOV-1996; 96BP-00308264.
XX
PR 15-NOV-1995; 95US-0006854P.
XX
PS (GENP-) GEN-PROBE INC.
PA Gordon P, Brentano ST, Carter NM, Hammond PW;
PI
DR WPI; 1997-274349/25.
XX
PT Probes for detection of Human Papillomavirus Type 16 and Type 18 - can
PT distinguish between Type 16 and 18, associated with genital cancers.
XX
PS Claim 16; Page 61; 70pp; English.
XX
CC Novel hybridisation assay probes have been developed comprising an
CC oligonucleotide which will hybridise under selected conditions to Human
CC Papillomavirus (HPV) Type 16 and/or 18 (but not Types 6, 11, 31, 33, 35,
CC 39, 45, 51, 52, or 58) target nucleic acids to form detectable target;
CC probe duplex. The present sequence represents a specifically claimed
CC target sequence. Oligonucleotides are useful to detect HPV Type 16 and/
CC or 18 in samples e.g. cervical smears, body fluid, and distinguish these
CC from other HPV variants. Papillomaviruses are small DNA viruses and HPV
CC Types 16 and 18 are associated with genital cancers. HPV PCR primers can
CC amplify HPV Type 16 and/or 18 nucleic acid in a sample. HPV Type 16
CC and/or 18 can be detected by adding a probe and detecting probe:target
CC duplex formation; target nucleic acid is optionally amplified. Type 16 or
CC 18 can be specifically detected by amplifying nucleic acids with at least
CC one specifically claimed PCR primer. For Type 18 detection, a helper
CC probe may be used. Oligonucleotides can distinguish target sequences from
CC phylogenetically close HPV Types 16 and 18 whilst available detection
CC sera react with antigens of all papillomaviruses
XX
SQ Sequence 32 BP; 8 A; 8 C; 7 G; 0 T; 9 U; 0 Other;

Query Match 100.0%; Score 32; DB 2; Length 32;
Best Local Similarity 100.0%; Pred. No. 9.2e-05;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DQ 1 TTATTAATAGTGCTGCGTGCCAGAAACC 32
32 TTATTAATAGTGCTGCGTGCCAGAAACC 1

```

RESULT 5
AAQ97848
ID AAQ97848 standard; cDNA; 476 BP.
XX
AC AAQ97848;
XX
DT 25-MAR-2003 (revised)
XX
DT 06-DEC-1995 (first entry)
XX
DR HPV-18 B6 cDNA.
XX
KM Ubiquitin-conjugating enzyme; HPV-18 B6 protein; cell cycle;
KM cell proliferation; cancer; psoriasis; fibrosis; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..477
FT /tag= a
XX
XX MO9518974-A2.
XX
PD 13-JUL-1995.
XX
PF 04-JAN-1995; 95MO-US000164.
XX
XX 04-JAN-1994; 94US-00176937.
PR 23-MAY-1994; 94US-00247904.
PR 27-MAY-1994; 94US-00250795.
PR 13-SEP-1994; 94US-00305520.
XX
PA (MITO-) MITOTIX INC.
PI Draetta G, Rolfe M, Eckstein JM, Cottarel G, Gyuris J;
DR WPI; 1995-255137/33.
DR P-PSDB; AAR9636.
XX
XX Identifying inhibitors of ubiquitin mediated proteolysis of cell cycle
XX regulatory proteins - also new ubiquitin conjugating enzymes, their
XX related nucleic acid, vectors, antibodies etc., useful for regulating
XX e.g. cell proliferation.
XX
XX Disclosure; Page 100-101; 157p; English.
XX
XX HPV-18 B6 cDNA (given in AAQ97848) was amplified from a HeLa cell cDNA
XX library using the primers given in AAQ97846-47. The gene was subcloned
XX into a baculovirus vector for expression of recombinant B6 in Sf9 insect
XX cells for use as a component of an in vitro ubiquitin conjugating system.
XX (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 476 BP; 163 A; 95 C; 100 G; 118 T; 0 U; 0 Other;
Query Match 100.0%; Score 32; DB 2; Length 476;
Best Local Similarity 100.0%; Pred. No. 0.00016;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TTATTAATAAGTGGCTGCGGTCCAGAAACC 32
Db 301 TTATTAATAAGTGGCTGCGGTCCAGAAACC 332
RESULT 6
AAZ27568
ID AAZ27568 standard; cDNA; 476 BP.
XX
AC AAZ27568;
XX
DT 15-DEC-1999 (first entry)
XX
DE HPV-18 B6 coding sequence.
XX
KM Ubiquitin conjugating enzyme; UbCE; ubiquitin-mediated proteolysis;

```

```

KM cell-cycle regulatory protein; ubiquitination inhibitor; atherosclerosis;
KM proliferative disorder; cancer; restenosis; tissue connective disorder;
KM wound healing; fibrosis; cancer; rheumatoid arthritis; scleroderma;
KM insulin dependent diabetes mellitus; glomerulonephritis; cirrhosis;
KM diagnosis; therapy; B6; ds.
XX
XX Human papillomavirus.
XX
XX US9568761-A.
XX
PD 19-OCT-1999.
XX
PF 07-JUN-1995; 95US-00486663.
XX
XX 04-JAN-1994; 94US-00176937.
PR 23-MAY-1994; 94US-00247904.
PR 27-MAY-1994; 94US-00250795.
PR 13-SEP-1994; 94US-00305520.
XX
XX (MITO-) MITOTIX INC.
XX
PI Chiu MI, Cottarel G, Berlin V, Damagnez V, Draetta G, Rolfe M;
XX WPI; 1999-590402/50.
XX
DR P-PSDB; AAY39968.
XX
XX Identifying ubiquitination inhibitors using novel ubiquitin conjugating
XX enzymes.
XX
XX Example 2; Col 89-92; 61pp; English.
XX
XX This sequence encodes the human papillomavirus B6 protein. The invention
XX relates to assays for identifying an inhibitor of ubiquitin-mediated
XX proteolysis of a cell-cycle regulatory protein comprising contacting a
XX candidate agent with an ubiquitin-conjugating system and measuring the
XX level of ubiquitination. The ubiquitin-conjugating system comprises: (a)
XX a reconstituted protein mixture including a ubiquitin conjugating enzyme
XX (UbCE) produced by the expression of a nucleic acid which hybridizes
XX under high stringency conditions to human UbCE, Candida albicans UbCE, or
XX Schizosaccharomyces pombe UbCE coding sequences; (b) a regulatory protein
XX ; and (c) ubiquitin. The polynucleotides are useful for identifying
XX ubiquitination inhibitors. The polynucleotides are useful for identifying
XX compounds and antibodies against them may also be useful for the
XX treatment and/or diagnosis of proliferative disorders (e.g. cancer,
XX atherosclerosis, or restenosis), tissue connective disorders, controlling
XX wound healing, and disorders characterized by fibrosis (e.g. rheumatoid
XX arthritis, insulin dependent diabetes mellitus, glomerulonephritis,
XX cirrhosis, and scleroderma)
XX
SQ Sequence 476 BP; 163 A; 95 C; 100 G; 118 T; 0 U; 0 Other;
Query Match 100.0%; Score 32; DB 2; Length 476;
Best Local Similarity 100.0%; Pred. No. 0.00016;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TTATTAATAAGTGGCTGCGGTCCAGAAACC 32
Db 301 TTATTAATAAGTGGCTGCGGTCCAGAAACC 332
RESULT 7
AAA61623
ID AAA61623 standard; cDNA; 476 BP.
XX
AC AAA61623;
XX
DT 15-SEP-2003 (revised)
XX
DT 23-OCT-2000 (first entry)
XX
DE cDNA encoding HPV-18 B6 protein.
XX
KM HPV-18 B6; ubiquitin mediated proteolysis; cellular protein half life;
KM ubiquitination inhibitor; p53; cyclin; cell cycle regulator;

```

myc deregulation; human papillomavirus; HPV-18 E6 protein;
KM cervical cancer; skin cancer; epidermal hyperplasia; epidermal neoplasia;
KM psoriasis; connective tissue disorder; wound healing; cytostatic;
KM antiproliferative; anticancer; antiposrotatic; ss.
XX
OS Human papillomavirus; 18.
XX
PH Key Location/Qualifiers
FT 1. .474
FT /*tag= a
FT /partial
FT /product= "HPV-18 E6 protein"
FT /note= "No stop codon given in the specification"
XX
XX US6068982-A.
XX
XX 30-MAY-2000.
XX
XX 17-DEC-1996; 96US-00767942.
XX
XX 04-JAN-1994; 94US-00176937.
XX 23-MAY-1994; 94US-00247904.
XX 27-MAY-1994; 94US-00250795.
XX 13-SEP-1994; 94US-00305520.
XX 07-JUN-1995; 95US-00486663.
XX
XX (MITO-) MITOTIX INC.
XX
XX Chiu MI, Cottarel G, Berlin V, Draetta G, Damagnez V, Rolfe M,
XX WPI; 2000-410854/35.
XX P-PSDB; AAB03176.
XX
XX Identifying an inhibitor of ubiquitin mediated proteolysis of regulatory
XX protein for treating cancers involves measuring ubiquitination levels of
XX the protein in the presence of candidate agent in an eukaryotic cell.
XX
XX Example 2; Col 97-98; 73pp; English.
XX
XX The invention relates to a method of identifying an inhibitor of
XX ubiquitin-mediated proteolysis of a cell cycle regulatory protein
XX comprising contacting an engineered eukaryotic cell with a candidate
XX agent. The eukaryotic cells is engineered to express a recombinant human,
XX Candida albicans or Schizosaccharomyces pombe ubiquitin-conjugating
XX enzyme (AAB03169-B03171), a cell cycle regulatory protein (such as p53)
XX and ubiquitin. The specification also discloses novel Candida albicans
XX and Schizosaccharomyces pombe ubiquitin-conjugating enzymes, caubCG and
XX spubCG (AAB03170, AAB03171), and two novel human ubiquitin-conjugating
XX enzymes, hubCG and rapubCG (AAB03169, AAB03173). The ubiquitin-mediated
XX proteolysis system is the major pathway for the selective, controlled
XX degradation of intracellular proteins in eukaryotic cells. In particular,
XX this system controls the half-lives of cellular proteins, and is
XX important in controlling the levels of proteins involved in cell cycle
XX progression. Alterations in the ubiquitination of these proteins may
XX therefore play a role in the development of cancers. For example, human
XX papillomaviruses such as HPV-18 encode a transforming protein, E6
XX (AAB03176), which combines with a cellular E6-associated protein (E6-AP,
XX AAB03177) to stimulate the ubiquitination of p53, thus targeting it for
XX degradation. The ubiquitination inhibitors identified according to the
XX method of the invention are useful for treatment of cervical cancers and
XX connective tissue disorders and for controlling the wound healing
XX process. They are also useful in treatment of hyperplastic epidermal
XX conditions such as psoriasis, neoplastic epidermal conditions, skin
XX cancers e.g., basal cell carcinomas, squamous cell carcinomas. The
XX inhibitors are useful for deregulating myc expression and rendering the
XX cells sensitive to chemotherapeutic treatment or to upset the balance of
XX transformed cells and cause apoptosis to occur. Inhibitors of ubiquitin-
XX mediated degradation of cyclins are useful as antiproliferative agents.
XX The present sequence represents cDNA encoding HPV-18 E6 protein. Note:
XX is given as SEQ ID NO:18 in the sequence listing. (Updated on 15-SEP-2003
XX to standardise OS field)

Sequence 476 BP; 163 A; 95 C; 100 G; 118 T; 0 U; 0 Other;
SQ
Query Match 100.0%; Score 32; DB 3; Length 476;
Best Local Similarity 100.0%; Pred. No. 0.00016;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TTATTATAAGTGGCTGCGTGCAGAAAC 32
DB 301 TTATTATAAGTGGCTGCGTGCAGAAAC 332
RESULT 8
AAL54443
ID AAL54443 standard; DNA; 477 BP.
XX
XX AAL54443;
XX
XX 03-APR-2003 (first entry)
XX
XX HPV18 E6 siRNA derived DNA sequence.
XX
XX Virucide; cytostatic; anti-HIV; dermatological; small interfering RNA;
XX selective post-translational silencing; siRNA; oncogene; genital wart;
XX human papilloma virus; HPV gene; cancer; human cervical cancer; HIV;
XX smallpox; flu; common cold; cervical cancer; penile cancer;
XX malignant squamous cell carcinoma; verruca vulgaris; gene therapy; ds.
XX
XX Human papilloma virus.
XX
XX WO2003068573-A2.
XX
XX 30-JAN-2003.
XX
XX 17-JUL-2002; 2002WO-GB003300.
XX
XX 17-JUL-2001; 2001GB-00017358.
XX 14-JUN-2002; 2002GB-00006688.
XX 17-JUN-2002; 2002GB-00013855.
XX
XX (MILN/) MILNER A J.
XX
XX Milner AJ;
XX
XX WPI; 2003-221850/21.
XX
XX Selective post-transcriptional silencing of an exogenous viral gene (e.g.
XX human papilloma virus (HPV) B6), for treating e.g. cancer, comprises
XX using a small interfering RNA (siRNA) construct homologous to an mRNA of
XX the gene.
XX
XX Claim 11; Fig 11; 44pp; English.
XX
XX The invention relates to a novel method for selective post-translational
XX silencing in a mammalian cell of the expression of an exogenous gene of
XX viral origin. The method comprises introducing into the cell a small
XX interfering RNA (siRNA) construct that is homologous to a part of the
XX mRNA sequence of the gene. The method is useful for the selective post-
XX transcriptional silencing of an exogenous gene of viral origin (e.g. an
XX oncogene or human papilloma virus (HPV) gene) in a mammalian cell. The
XX method or the siRNA is particularly useful for treating cancer, human
XX cervical cancer, human immunodeficiency virus (HIV), smallpox, flu,
XX common cold, or a disease caused by a HPV (e.g. genital warts, cervical
XX cancer, penile cancer, malignant squamous cell carcinomas or verruca
XX vulgaris). An siRNA construct or vector is useful for use as a medicament
XX for the diseases mentioned. The polynucleotide sequence of the invention
XX can be used to treat disorders by gene therapy. This polynucleotide
XX sequence represents the DNA of a HPV siRNA sequence of the invention
XX
XX Sequence 477 BP; 163 A; 96 C; 100 G; 118 T; 0 U; 0 Other;
SQ
Query Match 100.0%; Score 32; DB 10; Length 477;
Best Local Similarity 100.0%; Pred. No. 0.00016;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TTATTAATAAGTGGCTGCGTCCAGAAACC 32
 DB 301 TTATTAATAAGTGGCTGCGTCCAGAAACC 332

RESULT 9
 ADO44101
 ID ADO44101 standard; DNA; 477 BP.
 AC ADO44101;

XX 15-JUL-2004 (first entry)

DE Nucleotide sequence of the E6 polypeptide of HPV18.

XX E6 protein; E7 protein; fusion protein; HPV16; HPV-associated cancer;
 KM cervical cancer; immune response; lower gastrointestinal tract cancer;
 KM anal cancer; reproductive system cancer; penile cancer; vulvar cancer;
 KM gene; ss.

XX Human papillomavirus type 18.

XX Key Location/Qualifiers
 FT CDS 1..477
 FT /*tag= a
 FT /product= "E6 polypeptide"

PN MO2004030636-A2.

PD 15-APR-2004.

PF 02-OCT-2003; 2003MO-US031726.

PR 03-OCT-2002; 2002US-0415929P.

XX (AMHP) WYETH HOLDINGS CORP.

PI Smith L, Cassetti MC;

DR WPI; 2004-316328/29.

DR P-PSDB; ADO44074.

PT New polypeptide comprising human papillomavirus E6 and E7 polypeptides,
 useful for treating or preventing human papillomavirus (HPV)-associated
 PT cancers, e.g. cervical cancer.

XX Disclosure, Page 94; 101pp; English.

XX The present sequence encodes an E6 polypeptides from human papillomavirus
 CC type 18 (HPV18). The specification describes human papillomavirus E6 and
 CC E7 polypeptides, where the E7 polypeptide has mutations at any one or
 CC more of the amino acids corresponding to amino acids 24, 26 or 91 of the
 CC sequence given in ADO44073 and the E6 polypeptide has no mutations or has
 CC mutations at any one or more of the amino acids corresponding to amino
 CC acids 63 or 106 of the sequence given in ADO44072. The polypeptides of
 CC the invention are useful for treating or preventing human papillomavirus
 CC (HPV)-associated cancers, such as cervical cancer. The fusion proteins
 CC and nucleic acids encoding the fusion proteins are useful for generating
 CC immune responses against HPV. They are also useful for treating lower
 CC gastrointestinal tract cancers, e.g. anal cancer, and other cancers of
 CC the reproductive system, including penile and vulvar cancer.

XX Sequence 477 BP; 164 A; 96 C; 99 G; 118 T; 0 U; 0 Other;

Query Match 100.0%; Score 32; DB 12; Length 477;
 Best Local Similarity 100.0%; Pred. No. 0.00016;
 Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TTATTAATAAGTGGCTGCGTCCAGAAACC 32
 DB 301 TTATTAATAAGTGGCTGCGTCCAGAAACC 332

RESULT 10
 ID AEA51111
 AC AEA51111;
 DT 11-AUG-2005 (first entry)

XX Human papillomavirus type 18 E6/E7 gene fragment. SEQ ID NO: 15.

XX Delivery mechanism; gene therapy; cytostatic; vulnery; virucide;
 KM injury; infection; cancer; gastrointestinal disease;
 KM gynecology and obstetrics; tumor; colorectal tumor;
 KM uterine cervix tumor; squamous cell carcinoma; neoplasia; E7 gene; ds;
 KM E6 gene.

XX Human papillomavirus.

PN MO2005051431-A1.

XX 09-JUN-2005.

PF 25-NOV-2004; 2004MO-GB004979.

PR 25-NOV-2003; 2003GB-00027409.

PR 05-MAR-2004; 2004US-0549919P.

XX (MILN/) MILNER A J.

PI Milner AJ;

DR WPI; 2005-405310/41.

PT Composition useful for delivering an agent into a cell comprises the
 PT agent, a transfer agent and a solid or colloidal carrier medium.

XX Disclosure; SEQ ID NO 15; 56pp; English.

XX The present invention relates to a method and composition comprising a
 CC transfer agent and a solid or colloidal carrier medium for delivering a
 CC biological agents into cells. The invention is useful for the treatment
 CC of cancer, tumor, carcinoma of cutaneous, squamous or cervical epithelia
 CC and colorectal carcinoma, wounds, burns and scars. The invention is also
 CC useful in gene therapy. The present sequence is the human papillomavirus
 CC type 18 E6/E7 gene fragment. This sequence is useful in a method for
 CC treating carcinomas.

XX Sequence 477 BP; 164 A; 95 C; 100 G; 118 T; 0 U; 0 Other;

Query Match 100.0%; Score 32; DB 14; Length 477;
 Best Local Similarity 100.0%; Pred. No. 0.00016;
 Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TTATTAATAAGTGGCTGCGTCCAGAAACC 32
 DB 301 TTATTAATAAGTGGCTGCGTCCAGAAACC 332

RESULT 11
 ID AAO75471
 AC AAO75471;
 DT 16-OCT-2003 (revised)

DT 25-MAR-2003 (revised)

DT 28-JUN-1995 (first entry)

XX HPV18 E6/E7 proteins.

DE HPV18 E6/E7 proteins.

XX HPV18 E6 protein; E7 protein; diagnosis; cervical dysplasia;

KM cervix cancer; ss.

```

XX OS Human papillomavirus; strain 18.
XX FH Key Location/Qualifiers
XX FT CDS 1..480
XX FT /*tag= a
XX PN WO9426934-A2.
XX PD 24-NOV-1994.
XX PF 06-MAY-1994; 94MO-US005085.
XX PR 06-MAY-1993; 93US-00058920.
XX PA (BAKT ) BAXTER DIAGNOSTICS INC.
XX PI Brown JT;
XX DR WPI; 1995-006821/01.
XX DR P-PSDB; AAR63866.
XX PT Human papilloma virus detection assay - by amplification using self
XX PT sustained sequence replication and hybridisation with a detector probe.
XX PS Disclosure; Page 27-28; 79pp; English.
XX CC The sequences of the E6 and E7 polypeptide-encoding regions of human
XX CC papillomavirus (HPV) 16 and 18 are given in AAQ75470-71 and the encoded
XX CC proteins in AAR63865-66, respectively. Probes and primers based on these
XX CC sequences were used for HPV infection diagnosis; expression of E6 and E7
XX CC is diagnostic for cervical cancer or pre-malignant states. (Updated on
XX CC 25-MAR-2003 to correct PN field.) (Updated on 16-OCT-2003 to standardise
XX CC OS field)
XX SQ Sequence 483 BP; 167 A; 95 C; 100 G; 121 T; 0 U; 0 Other;
XX
Query Match 100.0%; Score 32; DB 2; Length 483;
Best Local Similarity 100.0%; Pred. No. 0.00016;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 TTATTATAAGTGCCTGCGTGCAGAAAC 32
Db 301 TTATTATAAGTGCCTGCGTGCAGAAAC 332

```

```

PI George AL, Groff DE;
XX DR WPI; 1989-324314/44.
XX PT Rapid detection of specific human papilloma virus genotypes - by
XX PT hybridisation of DNA digest with new labelled nucleic acid probes.
XX PS Claim 46; Page 48; 81pp; English.
XX CC Obt'd. by cutting HPV18 with BamH and Sau3A. The patent describes probes
XX CC (DNA or RNA) and their complements capable of detecting one or a
XX CC combination of HPV types 6, 11, 16, 18, 31, 33 and 35. (Updated on 25-MAR
XX CC -2003 to correct PR field.) (Updated on 27-AUG-2003 to correct OS field.)
XX SQ Sequence 802 BP; 269 A; 158 C; 182 G; 193 T; 0 U; 0 Other;
XX
Query Match 100.0%; Score 32; DB 1; Length 802;
Best Local Similarity 100.0%; Pred. No. 0.00018;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 TTATTATAAGTGCCTGCGTGCAGAAAC 32
Db 287 TTATTATAAGTGCCTGCGTGCAGAAAC 318

```

RESULT 13

```

AAQ29390
ID AAQ29390 standard; DNA; 817 BP.
XX AC AAQ29390;
XX DT 24-OCT-2003 (revised)
XX DT 25-MAR-2003 (revised)
XX DT 09-MAR-1993 (first entry)
XX DE DNA encoding HPV 18 E6/E7 proteins obt'd. by PCR.
XX KW Virus vector; vaccinia virus; papillomavirus; HPV; human; amplification;
XX KW immunotherapeutic; ss.
XX OS Human papillomavirus; 18.
XX FH Key Location/Qualifiers
XX FT CDS 1..816
XX FT /*tag= a
XX FT /note= "reading frame 1"
XX FT 2..815
XX FT /*tag= b
XX FT /note= "reading frame 2"
XX FT 3..814
XX FT /*tag= d
XX FT /note= "reading frame 3"
XX FT 6..482
XX FT /*tag= e
XX FT /note= "third reading frame encoding HPV 18 E6"
XX FT 491..808
XX FT /*tag= c
XX FT /note= "second reading frame encoding HPV 18 E7"
XX PN WO921636-A1.
XX PD 01-OCT-1992.
XX PF 10-MAR-1992; 92MO-GB000424.
XX PR 14-MAR-1991; 91GB-00005383.
XX PA (IMMU ) IMMUNOLOGY LTD.
XX XX Bourneell MEG, Inglis SC, Munro AJ;
XX PT WPI; 1992-349219/42.
XX DR P-PSDB; AAR27726, AAR27727, AAR27728.

```

XX Recombinant virus vectors encoding human papillomavirus proteins - for
 PT treating and vaccinating against HPV infections and conditions caused by
 PT chem, such as cervical cancer.
 XX Disclosure, Fig 1b; 83pp; English.
 CC The fragment of DNA contg. the HPV-18 E6/E7 coding region was prep'd. by
 CC PCR from plasmid pBR322/HPV18 (Boshart et al., EMBO J. 3: 1151) using
 CC oligonucleotide S01 and S02. The DNA prod. has a site for NcoI at the
 CC beginning of the E6 gene and a SmaI site immediately downstream of the
 CC termination codon for E7. The E6 and E7 ORFs are fused together to form a
 CC single continuous ORF via site directed mutagenesis and the immortalising
 CC potential of E7 is removed by altering two key codons of the HPV E7
 CC sequence. The single ORF of HPV-18 E6/E7 may be inserted into vaccinia
 CC virus DNA at neutral sites (pref. by inserting two sets of the DNA in
 CC opposite orientations to overcome the problem of intertypic
 CC recombination) to make a recombinant virus vector for use
 CC immunotherapeutically to activate cells of the immune system against HPV.
 CC See also AAO29385-400 and AAO29450-69. (Updated on 25-MAR-2003 to correct
 CC P1 field.) (Updated on 24-OCT-2003 to standardise OS field)
 CC
 SQ Sequence 817 BP; 266 A; 173 C; 177 G; 201 T; 0 U; 0 Other;
 Query Match 100.0%; Score 32; DB 2; Length 817;
 Best Local Similarity 100.0%; Pred. No. 0.00018;
 Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TTATTAATAGGTCCTGCGTCCAGAAACC 32
 Db 306 TTATTAATAGGTCCTGCGTCCAGAAACC 337
 RESULT 14
 ID AAX78800 standard; DNA; 837 BP.
 AC AAX78800;
 XX
 DT 06-SEP-1999 (first entry)
 XX
 DE HPV fusion protein D1/3-E6-His/HPV18 DNA.
 XX
 KM Fusion protein; E6 protein; E7 protein; E6/E7; immunomodulator; tumour;
 KM immunological fusion partner; CpG oligonucleotide; immune response;
 KM HPV antigen; prevention; treatment; ss.
 XX
 OS Synthetic.
 OS Human papillomavirus.
 XX
 PN WO9933868-A2.
 XX
 PD 08-JUL-1999.
 XX
 PF 18-DEC-1998; 98WO-EP008563.
 XX
 PR 24-DEC-1997; 97GB-00027262.
 XX
 PA (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
 PI Dalemans WJ, Gerard CMG;
 XX
 DR WPI; 1999-405485/34.
 DR P-PSDB; AAY25385.
 XX
 PT Composition comprising an E6, E7 or E6/E7 fusion protein from HPV to
 PT induce immune response to HPV.
 XX
 PS Example XI; Page 59; 62pp; English.
 XX
 CC AAX78791-X78801 represent nucleic acid sequences which encode novel
 CC constructs comprising an E6 or E7 protein or E6/E7 fusion protein from
 CC HPV (represented in AAY25375-Y25386). These constructs are optionally

CC linked to an immunological fusion partner and an immunomodulatory CpG
 CC oligonucleotide. The products of the invention can be used to induce an
 CC immune response in a patient to an HPV antigen. They can also be used for
 CC preventing or treating HPV induced tumours
 CC
 SQ Sequence 837 BP; 283 A; 169 C; 167 G; 218 T; 0 U; 0 Other;
 Query Match 100.0%; Score 32; DB 2; Length 837;
 Best Local Similarity 100.0%; Pred. No. 0.00018;
 Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TTATTAATAGGTCCTGCGTCCAGAAACC 32
 Db 634 TTATTAATAGGTCCTGCGTCCAGAAACC 665
 RESULT 15
 ID AAX29789 standard; DNA; 837 BP.
 AC AAX29789;
 XX
 DT 17-OCT-2003 (revised)
 DT 22-JUN-1999 (first entry)
 XX
 DE Prot.D1/3-E6-His/HPV18 coding sequence.
 XX
 KM Chimeric; E6; E7; fusion protein; protein D; vaccine; immunotherapy;
 KM tumour; lesion; benign; malignant; virus; infection; ss.
 XX
 OS Human papillomavirus.
 OS Haemophilus influenzae.
 OS Chimeric.
 XX
 PN WO9910375-A2.
 XX
 PD 04-MAR-1999.
 XX
 PF 17-AUG-1998; 98WO-BP005285.
 XX
 PR 22-AUG-1997; 97GB-00017953.
 XX
 PA (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
 PI Bruck C, Cabezon Silva T, Delisse ABF, Gerard CMG;
 PI Lombardo-Bencheikh A;
 XX
 DR WPI; 1999-190587/16.
 DR P-PSDB; AAY02641.
 XX
 PT Human Papilloma Virus (HPV) fusion proteins - useful in vaccines for
 PT treatment or prophylaxis of HPV induced lesions.
 XX
 PS Disclosure, Fig 22; 95pp; English.
 XX
 CC This sequence represents the coding region for a chimeric E6 or E7
 CC protein or E6/E7 fusion protein from human papillomavirus (HPV) linked to
 CC an immunological fusion partner. In this case, a fragment of the
 CC Haemophilus influenzae B protein D. The sequence also contains a
 CC histidine tag at the C-terminus of the encoded protein. The protein can
 CC be used in a vaccine, for immuno-therapeutically treating HPV induced
 CC tumour lesions (benign or malignant) and preventing HPV viral infection.
 CC (Updated on 17-OCT-2003 to standardise OS field)
 CC
 SQ Sequence 837 BP; 283 A; 169 C; 167 G; 218 T; 0 U; 0 Other;
 Query Match 100.0%; Score 32; DB 2; Length 837;
 Best Local Similarity 100.0%; Pred. No. 0.00018;
 Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TTATTAATAGGTCCTGCGTCCAGAAACC 32
 Db 634 TTATTAATAGGTCCTGCGTCCAGAAACC 665


```
RESULT 16
AED52652
ID AED52652 standard; DNA; 837 BP.
XX
XX
AC AED52652;
XX
DT 29-DEC-2005 (first entry)
XX
XX Fusion protein D1/3-E6-His (HPV18) , DNA.
XX
XX Fusion protein; vaccine; papilloma; cyostatic; papillomavirus infection;
XX virucide; uterine cervix tumor; E6; ds; gene; D protein.
XX
XX Haemophilus influenzae; strain 772.
XX
XX Human papillomavirus type 18.
XX
XX Synthetic.
XX
XX Chimeric.
XX
XX Key Location/Qualifiers
XX CDS 1..837
XX /*tag= a
XX /product= "Fusion protein D1/3-E6-His (HPV18) "
XX
XX IN9801903-14.
XX
XX 04-MAR-2005.
XX
XX 24-AUG-1998; 98IN-CH001903.
XX
XX 22-AUG-1997; 97EP-00179535.
XX
XX (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.
XX
XX Tyrrell AMR;
XX
XX WPI; 2005-557648/57.
XX
XX P-PSDB; AED52653.
XX
XX Vaccine.
XX
XX Example 16; Fig 22; 96pp; English.
XX
XX The invention relates to human Papilloma virus (HPV) fusion proteins,
XX linked to an immunological fusion partner that provides T helper epitopes
XX to the HPV antigen. Vaccine formulations comprising the fusion proteins
XX are useful in the treatment or prophylaxis of HPV induced lesions
XX (papillomas, warts and cervical cancer). HPV 16 (and HPV 18) E6 and E7
XX proteins (singly, as an E6-E7 fusion or mutated) were fused to either
XX Haemophilus influenzae D protein (20-127), the C-termius of
XX Streptococcus pneumoniae Lyta protein (cLYTA) or chlorodoxin. The present
XX sequence encodes an HPV-H. influenzae D protein, fusion protein of the
XX invention.
XX
XX Sequence 837 BP; 283 A; 169 C; 167 G; 218 T; 0 U; 0 Other;
SQ
Query Match 100.0%; Score 32; DB 14; Length 837;
Best Local Similarity 100.0%; Pred. No. 0.00018;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TTATTATTAAGGTGCTGCGGTGCCAGAAACC 32
DB 634 TTATTATTAAGGTGCTGCGGTGCCAGAAACC 665
```

```
XX
DE Human papillomavirus E6 and E7 coding sequence, SEQ ID 29.
XX
XX Cytostatic; Virucide; Antibacterial; Antiparasitic; Vaccine; infection;
XX cancer; gene; ds; antigen.
XX
XX Human papillomavirus.
XX
XX Key Location/Qualifiers
XX CDS 40..516
XX /*tag= a
XX /product= "E6 protein"
XX CDS 525..842
XX /*tag= b
XX /product= "E7 protein"
XX
XX US2004258688-A1.
XX
XX 23-DEC-2004.
XX
XX 12-MAR-2004; 2004US-00800023.
XX
XX 31-JAN-1995; 95US-00381528.
XX
XX 31-JAN-1996; 96WO-US001383.
XX
XX 05-JUN-2000; 2000US-00586704.
XX
XX 09-AUG-2001; 2001US-00925284.
XX
XX (HAWI/) HAWIGER D.
XX (NUSZ/) NUSSENZWEIG M.
XX (STEI/) STEINMAN R M.
XX (BONI/) BONIFAZ L.
XX
XX Hawiger D, Nussenzweig M, Steinman RM, Bonifaz L;
XX
XX WPI; 2005-078933/09.
XX
XX P-PSDB; ADV85646, ADV85647.
XX
XX Promoting antigen presentation comprises targeting antigen to dendritic
XX cells using an anti-DEC-205 antibody.
XX
XX Disclosure; SEQ ID NO 29; 116pp; English.
XX
XX The present invention relates to a method (M1) for promoting highly
XX efficient antigen presentation in a mammal, by targeting a preselected
XX antigen to an endocytic receptor on an antigen-presenting cell, e.g. a
XX dendritic cell. An example of a dendritic cell is the cell endocytic
XX receptor DEC-205. The method comprises exposing dendritic cells from the
XX mammal to either: a conjugate comprising a preselected antigen covalently
XX bound to an antibody; a conjugate comprising a preselected antigen on at
XX genetically engineered to contain at least one preselected antigen on at
XX least one preselected site on the antibody molecule; and promoting
XX maturation of the dendritic cells. The sequence encoding an anti-DEC-205
XX antibody (fragment) is preferably selected from ADV85632 and ADV85633,
XX which encode the heavy or light chain variable region of an anti-DEC-205
XX antibody. The antigen is targeted to antigen presenting cells through the
XX inclusion of the anti-DEC-205 antibody, making antigen presentation
XX highly efficient. The immunity induced is robust and long lasting, even
XX from a single dose at low concentration. The method of the invention is
XX useful for immunizing a mammal to prevent or treat a disease such as a
XX viral, bacterial or other infection or cancer. The immunization
XX preferably results in induction of long term T cell, B cell or mucosal
XX immunity. The present sequence is a HPV DNA sequence which encodes the E6
XX and E7 proteins. The present sequence can be used as a vaccine antigen.
XX
XX Sequence 843 BP; 283 A; 170 C; 183 G; 207 T; 0 U; 0 Other;
SQ
Query Match 100.0%; Score 32; DB 14; Length 843;
Best Local Similarity 100.0%; Pred. No. 0.00018;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TTATTATTAAGGTGCTGCGGTGCCAGAAACC 32
DB 340 TTATTATTAAGGTGCTGCGGTGCCAGAAACC 371
```

```

RESULT 18
AAQ08623
ID AAQ08623 standard; DNA; 1000 BP.
XX
AC AAQ08623;
XX
DT 24-OCT-2003 (revised)
DT 21-APR-1994 (first entry)
XX
DE HPV-18 fragment.
XX
XX HPV-18; HPV-16; amplification; primer; polymerase chain reaction; PCR;
KM 88.
XX
OS Human papillomavirus; 18.
XX
FH Key Location/Qualifiers
FT misc_binding 167..186
FT /*tag= a
FT /note= "primer (AAQ08624) binding site"
FT 647..666
FT /*tag= c
FT /note= "primer (AAQ08626) binding site"
FT 667..686
FT /*tag= b
FT /note= "primer (AAQ08625) binding site"
XX
PN DE3838269-A.
XX
PD 17-MAY-1990.
XX
PF 11-NOV-1988; 88DE-03838269.
XX
PR 11-NOV-1988; 88DE-03838269.
XX
PA (BEHW ) BEHRINGERWERKE AG.
XX
PI Cerutti P, Whitcomb J, Zijlstra J, Devillers EM;
XX WPI; 1990-156905/21.
XX
XX
XX Detection of human papilloma virus - by DNA amplification and analysis.
XX
PS Example 4a; Page 4; 11pp; German.
XX
CC Example 4a describes the results of the amplification of HPV-18 DNA by
CC PCR using primers. (Updated on 24-OCT-2003 to standardise OS field)
XX
SQ Sequence 1000 BP; 341 A; 198 C; 221 G; 240 T; 0 U; 0 Other;
XX
Query Match 100.0%; Score 32; DB 2; Length 1000;
Best Local Similarity 100.0%; Pred. No. 0.00019;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 TTATTATAAGTGCTGCGTGCCAGAAACC 32
Db 405 TTATTATAAGTGCTGCGTGCCAGAAACC 436

RESULT 19
AAK78801
ID AAK78801 standard; DNA; 1152 BP.
XX
AC AAK78801;
XX
DT 06-SEP-1999 (first entry)
XX
DE HPV fusion protein D1/3-E6/E7-His/HPV18 DNA.
XX
KM Fusion protein; E6 protein; E7 protein; E6/E7; immunomodulator; tumour;
KM immunological fusion partner; Cpg oligonucleotide; immune response;

```

```

KM HPV antigen; prevention; treatment; ss.
XX
OS Synthetic.
OS Human papillomavirus.
XX
PN W09933868-A2.
XX
PD 08-JUL-1999.
XX
PF 18-DEC-1998; 98MO-EP008563.
XX
PR 24-DEC-1997; 97GB-00027262.
XX
PA (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.
XX
PI Dalemans WLJ, Gerard CMG;
XX
XX WPI; 1999-405485/34.
DR P-PSDB; AAY25386.
XX
PT Composition comprising an E6, E7 or E6/E7 fusion protein from HPV to
PT induce immune response to HPV.
XX
PS Example XII; Page 60-61; 62pp; English.
XX
XX AAK78791-X78801 represent nucleic acid sequences which encode novel
CC constructs comprising an E6 or E7 protein or E6/E7 fusion protein from
CC HPV (represented in AAY25375-Y25386). These constructs are optionally
CC linked to an immunological fusion partner and an immunomodulatory Cpg
CC oligonucleotide. The products of the invention can be used to induce an
CC immune response in a patient to an HPV antigen. They can also be used for
CC preventing or treating HPV induced tumours
XX
SQ Sequence 1152 BP; 378 A; 240 C; 240 G; 294 T; 0 U; 0 Other;
XX
Query Match 100.0%; Score 32; DB 2; Length 1152;
Best Local Similarity 100.0%; Pred. No. 0.00019;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 TTATTATAAGTGCTGCGTGCCAGAAACC 32
Db 634 TTATTATAAGTGCTGCGTGCCAGAAACC 665

RESULT 20
AAK29790
ID AAK29790 standard; DNA; 1152 BP.
XX
AC AAK29790;
XX
DT 17-OCT-2003 (revised)
DT 22-JUN-1999 (first entry)
XX
DE Prot. D1/3-E6-E7-His/HPV18 coding sequence.
XX
KM Chimeric; E6; E7; fusion protein; protein D; vaccine; immunotherapy;
KM tumour; lesion; benign; malignant; virus; infection; ss.
XX
XX Human papillomavirus.
OS Haemophilus influenzae.
OS Chimeric.
XX
XX W09910375-A2.
XX
PD 04-MAR-1999.
XX
PF 17-AUG-1998; 98MO-EP005285.
XX
PR 22-AUG-1997; 97GB-00017953.
XX
PA (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.
XX
PI Bruck C, Caberon Silva T, Delisse ABF, Gerard CMG;

```

PI Lombardo-Bencheikh A;
XX WPI; 1999-190587/16.
DR P-PSDB; AAY02642.
XX Human Papilloma Virus (HPV) fusion proteins - useful in vaccines for
PT treatment or prophylaxis of HPV induced lesions.
XX
PS Disclosure; Fig 25; 95pp; English.
XX
CC This sequence represents the coding region for a chimeric E6 or E7
CC protein or E6/E7 fusion protein from Human papillomavirus (HPV) linked to
CC an immunological fusion partner, in this case, a fragment of the
CC Haemophilus influenzae B protein D. The sequence also contains a
CC histidine tag at the C-terminus of the encoded protein. The protein can
CC be used in a vaccine, for immuno-therapeutically treating HPV induced
CC tumour lesions (benign or malignant) and preventing HPV viral infection.
CC (Updated on 17-OCT-2003 to standardise OS field)
XX
SQ Sequence 1152 BP; 378 A; 240 C; 240 G; 294 T; 0 U; 0 Other;
Query Match 100.0%; Score 32; DB 2; Length 1152;
Best Local Similarity 100.0%; Pred. No. 0.00019;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TTATTAATAAGTGCCTGCGGTGCCAGAAACC 32
DB 634 TTATTAATAAGTGCCTGCGGTGCCAGAAACC 665
RESULT 21
AED52656
ID AED52656 standard; DNA; 1152 BP.
XX
AC AED52656;
XX
DT 29-DEC-2005 (first entry)
XX
DE Fusion protein D1/3-E6-E7-His/HPV18, DNA.
XX
KM Fusion protein; vaccine; papilloma; cytostatic; papillomavirus infection;
KM virulence; uterine cervix tumor; E7; E6; ds; gene; D protein.
XX
OS Haemophilus influenzae; strain 772.
OS Human papillomavirus type 18.
OS Synthetic.
OS Chimeric.
XX
FH Key Location/Qualifiers
FT CDS 1..1152
FT /*tag= a
FT /product= "Fusion protein D1/3-E6-E7-His/HPV18"
PN IN9801903-14.
XX
PD 04-MAR-2005.
XX
PE 24-AUG-1998; 98IN-CH001903.
XX
PR 22-AUG-1997; 97BP-00179535.
XX
PA (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
XX
PI Tyrell AMR;
XX
DR WPI; 2005-557648/57.
DR P-PSDB; AED52657.
XX
PT Vaccine.
XX
PS Example 17; Fig 25; 96pp; English.
XX
CC The invention relates to human Papilloma virus (HPV) fusion proteins,

CC linked to an immunological fusion partner that provides T helper epitopes
CC to the HPV antigen. Vaccine formulations comprising the fusion proteins
CC are useful in the treatment or prophylaxis of HPV induced lesions
CC (papillomas, warts and cervical cancer). HPV 16 (and HPV 18) E6 and E7
CC proteins (singly, as an E6-E7 fusion or mutated) were fused to either
CC Haemophilus influenzae D protein (20-127), the C-terminus of
CC Streptococcus pneumoniae Lyta protein (cLYTA) or thiorodoxin. The present
CC sequence encodes an HPV-H. influenzae D protein, fusion protein of the
CC invention.
XX
SQ Sequence 1152 BP; 378 A; 240 C; 240 G; 294 T; 0 U; 0 Other;
Query Match 100.0%; Score 32; DB 14; Length 1152;
Best Local Similarity 100.0%; Pred. No. 0.00019;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TTATTAATAAGTGCCTGCGGTGCCAGAAACC 32
DB 634 TTATTAATAAGTGCCTGCGGTGCCAGAAACC 665
RESULT 22
AAN80161
ID AAN80161 standard; DNA; 1731 BP.
XX
AC AAN80161;
XX
DT 27-AUG-2003 (revised)
DT 25-MAR-2003 (revised)
DT 10-MAR-2003 (revised)
DT 19-OCT-1990 (first entry)
XX
DE Human papilloma virus 18 protein.
XX
KM Human papilloma virus; HPV16; HPV18; malignant genital tumours; ss.
XX
OS Human papillomavirus.
XX
FH Key Location/Qualifiers
FT CDS 99..572
FT /*tag= a
FT /label= HPV18 E6 protein
FT 584..898
FT /*tag= b
FT /label= HPV 18 E7 protein
FT 908..729
FT /*tag= c
FT /label= HPV 18 B1 protein
XX
PN DE3625257-A.
XX
PD 04-FEB-1988.
XX
PE 25-JUL-1986; 86DE-03625257.
XX
PR 23-JUL-1986; 86DE-03624786.
XX
PA (BEHM) BEHRINGWERKE AG.
XX
PI Schwarz B, Oltersdorf T, Rowekamp W, Schneiderg A, Seedorf K;
PI Durst M;
XX
DR WPI; 1988-037095/06.
DR P-PSDB; AAP82076.
XX
PT Expression prods. of human papilloma viruses type 16 and 18 - used for
PT the prodn. of diagnostics for HPV16 and HPV18 infections.
XX
PS Disclosure; Page 7; 7pp; German.
XX
CC The expression prods. of HPV18, contained in this sequence, are used to
CC produce antibodies to detect these. HPV16 expression prods are also used
CC for antibody prodn. See also AAN80155-59. (Updated on 10-MAR-2003 to add

CC missing OS field.) (Updated on 25-MAR-2003 to correct PR field.) (Updated
 CC on 25-MAR-2003 to correct PI field.) (Updated on 27-AUG-2003 to correct
 CC OS field.)
 XX
 SO Sequence 1731 BP; 610 A; 313 C; 403 G; 405 T; 0 U; 0 Other;
 Query Match 100.0%; Score 32; DB 1; Length 1731;
 Best Local Similarity 100.0%; Pred. No. 0.00021;
 Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TTATTAATAGGCTGCGTCCGAGAAC 32
 DB 399 TTATTAATAGGCTGCGTCCGAGAAC 430
 RESULT 23
 AD245648
 ID AD245648 standard; DNA; 7857 BP.
 XX
 AC AD245648;
 XX
 DT 30-JUN-2005 (first entry)
 XX
 DE Human papillomavirus type 18 full length genomic clone - SEQ ID 2.
 XX
 KW tumor marker; genome; diagnosis; cancer; uterine cervix tumor;
 KM rectal tumor; colon tumor; ds.
 XX
 OS Human papillomavirus type 18.
 XX
 PN WO200503333-A2.
 XX
 PD 14-APR-2005.
 XX
 PF 04-OCT-2004; 2004WO-DK000670.
 XX
 PR 07-OCT-2003; 2003DK-00001474.
 PR 07-OCT-2003; 2003US-0509205P.
 PR 13-FEB-2004; 2004US-0543925P.
 XX
 PA (DAKO-) DAKOCYTOMATION DENMARK AS.
 XX
 PI Chau MF, Bisgaard-Franzen K, Lin J, Raasmussen OF, Wang Z, Lusk J,
 PI Lindberg M, Yeast S;
 XX
 DR WPI; 2005-285441/29.
 XX
 PT New composition having a nucleic acid molecule identical to genomic clone
 PT of human papilloma virus 16, 18 or 51, useful in diagnosing cancer or the
 PT risk of developing cancer, in particular cervical, anal, colon and HPV-
 PT related cancer.
 XX
 PS Claim 1; SEQ ID NO 2; 169pp; English.
 XX
 CC The invention comprises a composition for the detection of cancer
 CC markers, the composition contains a nucleic acid molecule which is
 CC substantially identical to a full length genomic clone of a human
 CC papilloma virus (type 11, 16, 18, 51, 56, 58, 66, 70, or 73). The
 CC composition of the invention is useful for diagnosing cancer or the risk
 CC of developing cancer, in particular cervical, anal, colon and HPV-related
 CC cancer. The present DNA sequence represents a human papilloma virus full
 CC length genomic clone of the invention.
 CC
 SO Sequence 7857 BP; 2365 A; 1497 C; 1680 G; 2315 T; 0 U; 0 Other;
 Query Match 100.0%; Score 32; DB 14; Length 7857;
 Best Local Similarity 100.0%; Pred. No. 0.00029;
 Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TTATTAATAGGCTGCGTCCGAGAAC 32
 DB 405 TTATTAATAGGCTGCGTCCGAGAAC 436

RESULT 24
 ADO44106
 ID ADO44106 standard; DNA; 477 BP.
 XX
 AC ADO44106;
 XX
 DT 15-JUL-2004 (first entry)
 XX
 DE Nucleotide sequence of the E6 polypeptide of HPV45.
 XX
 KW E6 protein; E7 protein; fusion protein; HPV45; HPV-associated cancer;
 KW cervical cancer; immune response; lower gastrointestinal tract cancer;
 KW anal cancer; reproductive system cancer; penile cancer; vulvar cancer;
 KW gene; ss.
 XX
 OS Human papillomavirus type 45.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..477
 FT /tag= a
 FT /product= "E6 polypeptide"
 XX
 PN WO2004030636-A2.
 XX
 PD 15-APR-2004.
 XX
 PF 02-OCT-2003; 2003WO-US031726.
 XX
 PR 03-OCT-2002; 2002US-0415929P.
 XX
 PA (AMHP) WYETH HOLDINGS CORP.
 XX
 PI Smith L, Cassetti MC;
 XX
 DR WPI; 2004-316328/29.
 DR P-PSDB; ADO44079.
 XX
 PT New polypeptide comprising human papillomavirus E6 and E7 polypeptides,
 PT useful for treating or preventing human papillomavirus (HPV)-associated
 PT cancers, e.g. cervical cancer.
 XX
 PS Disclosure; Page 96; 101pp; English.
 XX
 CC The present sequence encodes an E6 polypeptides from human papillomavirus
 CC type 45 (HPV45). The specification describes human papillomavirus E6 and
 CC E7 polypeptides, where the E7 polypeptide has mutations at any one or
 CC more of the amino acids corresponding to amino acids 24, 26 or 91 of the
 CC sequence given in ADO44073 and the E6 polypeptide has no mutations or has
 CC mutations at any one or more of the amino acids corresponding to amino
 CC acids 63 or 106 of the sequence given in ADO44072. The polypeptides of
 CC the invention are useful for treating or preventing human papillomavirus
 CC (HPV)-associated cancers, such as cervical cancer. The fusion proteins
 CC and nucleic acids encoding the fusion proteins are useful for generating
 CC immune responses against HPV. They are also useful for treating lower
 CC gastrointestinal tract cancers, e.g. anal cancer, and other cancers of
 CC the reproductive system, including penile and vulvar cancer.
 XX
 SO Sequence 477 BP; 171 A; 86 C; 101 G; 119 T; 0 U; 0 Other;
 Query Match 95.0%; Score 30.4; DB 12; Length 477;
 Best Local Similarity 96.9%; Pred. No. 0.00089;
 Matches 31; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 TTATTAATAGGCTGCGTCCGAGAAC 32
 DB 301 TTATTAATAGGCTGCGTCCGAGAAC 332
 RESULT 25
 ADO44105
 ID ADO44105 standard; DNA; 477 BP.
 XX

```

AC ADO44105;
XX
XX 15-JUL-2004 (first entry)
XX
XX Nucleotide sequence of the E6 polypeptide of HPV39.
XX
XX E6 protein; E7 protein; fusion protein; HPV39; HPV-associated cancer;
XX cervical cancer; immune response; lower gastrointestinal tract cancer;
XX anal cancer; reproductive system cancer; penile cancer; vulvar cancer;
XX gene; ss.
XX
XX Human papillomavirus type 39.
XX
XX
XX Key Location/Qualifiers
XX CDS 1..477
XX /*tag= a
XX /product= "E6 polypeptide"
XX
XX WO2004030636-A2.
XX
XX 15-APR-2004.
XX
XX 02-OCT-2003; 2003WO-US031726.
XX
XX 03-OCT-2002; 2002US-0415929P.
XX
XX (AMHP ) WYETH HOLDINGS CORP.
XX
XX Smith L, Cassetti MC;
XX
XX WPI; 2004-316328/29.
XX
XX P-PSDB; ADO44078.
XX
XX New polypeptide comprising human papillomavirus E6 and E7 polypeptides,
XX useful for treating or preventing human papillomavirus (HPV)-associated
XX cancers, e.g. cervical cancer.
XX
XX Disclosure; Page 95-96; 101pp; English.
XX
XX The present sequence encodes an E6 polypeptides from human papillomavirus
XX type 39 (HPV39). The specification describes human papillomavirus E6 and
XX E7 polypeptides, where the E7 polypeptide has mutations at any one or
XX more of the amino acids corresponding to amino acids 24, 26 or 91 of the
XX sequence given in ADO44073 and the E6 polypeptide has no mutations or has
XX mutations at any one or more of the amino acids corresponding to amino
XX acids 63 or 106 of the sequence given in ADO44072. The polypeptides of
XX the invention are useful for treating or preventing human papillomavirus
XX (HPV)-associated cancers, such as cervical cancer. The fusion proteins
XX and nucleic acids encoding the fusion proteins are useful for generating
XX immune responses against HPV. They are also useful for treating lower
XX gastrointestinal tract cancers, e.g. anal cancer, and other cancers of
XX the reproductive system, including penile and vulvar cancer.
XX
XX Sequence 477 BP; 170 A; 96 C; 100 G; 111 T; 0 U; 0 Other;
XX
XX Query Match 75.0%; Score 24; DB 12; Length 477;
XX Best Local Similarity 84.4%; Pred. No. 0.84;
XX Matches 27; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
XX
XX 1 TTATTATAAGGTCGCGTGCAGAAACC 32
XX Db TTATTATAAGGTCAGTGTGCTGAAACC 332
XX
XX
XX RESULT 26
XX ADO44112
XX ID ADO44112 standard; DNA; 477 BP.
XX
XX ADO44112;
XX
XX 15-JUL-2004 (first entry)
XX
XX Nucleotide sequence of the E6 polypeptide of HPV68.
XX

```

```

XX
XX E6 protein; E7 protein; fusion protein; HPV68; HPV-associated cancer;
XX cervical cancer; immune response; lower gastrointestinal tract cancer;
XX anal cancer; reproductive system cancer; penile cancer; vulvar cancer;
XX gene; ss.
XX
XX Human papillomavirus type 68.
XX
XX
XX Key Location/Qualifiers
XX CDS 1..477
XX /*tag= a
XX /product= "E6 polypeptide"
XX
XX WO2004030636-A2.
XX
XX 15-APR-2004.
XX
XX 02-OCT-2003; 2003WO-US031726.
XX
XX 03-OCT-2002; 2002US-0415929P.
XX
XX (AMHP ) WYETH HOLDINGS CORP.
XX
XX Smith L, Cassetti MC;
XX
XX WPI; 2004-316328/29.
XX
XX P-PSDB; ADO44085.
XX
XX New polypeptide comprising human papillomavirus E6 and E7 polypeptides,
XX useful for treating or preventing human papillomavirus (HPV)-associated
XX cancers, e.g. cervical cancer.
XX
XX Disclosure; Page 98; 101pp; English.
XX
XX The present sequence encodes an E6 polypeptides from human papillomavirus
XX type 68 (HPV68). The specification describes human papillomavirus E6 and
XX E7 polypeptides, where the E7 polypeptide has mutations at any one or
XX more of the amino acids corresponding to amino acids 24, 26 or 91 of the
XX sequence given in ADO44073 and the E6 polypeptide has no mutations or has
XX mutations at any one or more of the amino acids corresponding to amino
XX acids 63 or 106 of the sequence given in ADO44072. The polypeptides of
XX the invention are useful for treating or preventing human papillomavirus
XX (HPV)-associated cancers, such as cervical cancer. The fusion proteins
XX and nucleic acids encoding the fusion proteins are useful for generating
XX immune responses against HPV. They are also useful for treating lower
XX gastrointestinal tract cancers, e.g. anal cancer, and other cancers of
XX the reproductive system, including penile and vulvar cancer.
XX
XX Sequence 477 BP; 167 A; 93 C; 102 G; 115 T; 0 U; 0 Other;
XX
XX Query Match 75.0%; Score 24; DB 12; Length 477;
XX Best Local Similarity 84.4%; Pred. No. 0.84;
XX Matches 27; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
XX
XX 1 TTATTATAAGGTCGCGTGCAGAAACC 32
XX Db TTATTATAAGGTCAGTGTGCTGAAACC 332
XX
XX
XX RESULT 27
XX AA231781
XX ID AA231781 standard; DNA; 1095 BP.
XX
XX AA231781;
XX
XX 21-JAN-2000 (first entry)
XX
XX HPV68 DNA sequence.
XX
XX HPV68; human papillomavirus 68; detection; immunisation; infection; ss.
XX
XX Human papillomavirus.
XX

```

```
PN US5981173-A.
XX
XX 09-NOV-1999.
XX
PF 11-FEB-1997; 97US-00815667.
XX
XX 14-FEB-1996; 96US-0011650P.
PR 15-FEB-1996; 96US-0020458P.
XX
PA (INSP ) INST PASTEUR.
PA (INRM ) INST NAT SANTE & RECH MEDICALE.
XX
XX Longuet M, Orth G, Beaudenon S;
XX WPI; 1999-633305/54.
XX
XX New human papillomavirus (HPV) 68 and 70 DNA sequences, useful in the
XX treatment of HPV infections.
XX
XX Claim 2; Fig 5a; 25pp; English.
XX
CC This sequence represents a human papillomavirus (HPV) 68 DNA sequence.
CC The invention relates to a purified HPV68 DNA (I) deposited at C.N.C.M.
CC (Collection Nationale de Culture de Microorganismes) under Accession No. I
CC -1540, and a HPV70 DNA (II) sequence. Fragments of the HPV68 and HPV70
CC DNA may be useful as primers or probes in the detection of HPV68 or HPV70
CC infections. Proteins produced by the HPV DNA's can be used to immunise
CC animals against HPV68 or HPV70 infections
XX
SQ Sequence 1095 BP; 370 A; 220 C; 231 G; 274 T; 0 U; 0 Other;
XX
Query Match 75.0%; Score 24; DB 2; Length 1095;
Best Local Similarity 84.4%; Pred. No. 0.99;
Matches 27; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 1 TTATTAATAAGTGCTGCGGTCGCAAAACC 32
Db 580 TTATGATAGTGATGATGATGCTGCAAAACC 611
XX
RESULT 28
AAZ31783
ID AAZ31783 standard; DNA; 1108 BP.
XX
AC AAZ31783;
XX
XX 21-JAN-2000 (first entry)
XX
XX HPV68 E6 and E7 protein coding sequence.
XX
XX HPV68; human papillomavirus 68; detection; immunisation; infection;
XX E6 protein; E7 protein; ss.
XX
XX Human papillomavirus.
XX
XX US5981173-A.
XX
XX 09-NOV-1999.
XX
XX 11-FEB-1997; 97US-00815667.
XX
XX 14-FEB-1996; 96US-0011650P.
PR 15-FEB-1996; 96US-0020458P.
XX
XX (INSP ) INST PASTEUR.
XX (INRM ) INST NAT SANTE & RECH MEDICALE.
XX
XX Longuet M, Orth G, Beaudenon S;
XX WPI; 1999-633305/54.
XX
XX New human papillomavirus (HPV) 68 and 70 DNA sequences, useful in the
XX treatment of HPV infections.
XX
XX
```

```
XX
XX Claim 4; Fig 3a; 25pp; English.
XX
XX This sequence encodes the E6 and E7 proteins of human papillomavirus
XX (HPV) 68. The invention relates to a purified HPV68 DNA (I) deposited at
XX C.N.C.M. (Collection Nationale de Culture de Microorganismes) under
XX Accession No. I-1540, and a HPV70 DNA (II) sequence. Fragments of the
XX HPV68 and HPV70 DNA may be useful as primers or probes in the detection
XX of HPV68 or HPV70 infections. Proteins produced by the HPV DNA's can be
XX used to immunise animals against HPV68 or HPV70 infections
XX
XX Sequence 1108 BP; 375 A; 223 C; 231 G; 279 T; 0 U; 0 Other;
XX
Query Match 75.0%; Score 24; DB 2; Length 1108;
Best Local Similarity 84.4%; Pred. No. 1;
Matches 27; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 1 TTATTAATAAGTGCTGCGGTCGCAAAACC 32
Db 593 TTATGATAGTGATGATGATGCTGCAAAACC 624
XX
RESULT 29
AAZ31789
ID AAZ31789 standard; DNA; 3283 BP.
XX
XX AAZ31789;
XX
XX 21-JAN-2000 (first entry)
XX
XX HPV70 DNA sequence.
XX
XX HPV68; human papillomavirus 68; detection; immunisation; infection;
XX HPV70; ss.
XX
XX Human papillomavirus.
XX
XX US5981173-A.
XX
XX 09-NOV-1999.
XX
XX 11-FEB-1997; 97US-00815667.
XX
XX 14-FEB-1996; 96US-0011650P.
PR 15-FEB-1996; 96US-0020458P.
XX
XX (INSP ) INST PASTEUR.
XX (INRM ) INST NAT SANTE & RECH MEDICALE.
XX
XX Longuet M, Orth G, Beaudenon S;
XX WPI; 1999-633305/54.
XX
XX New human papillomavirus (HPV) 68 and 70 DNA sequences, useful in the
XX treatment of HPV infections.
XX
XX Disclosure; Fig 6; 25pp; English.
XX
XX This sequence represents a human papillomavirus (HPV) 70 DNA sequence.
XX The invention relates to a purified HPV68 DNA (I) deposited at C.N.C.M.
XX (Collection Nationale de Culture de Microorganismes) under Accession No. I
XX -1540, and a HPV70 DNA (II) sequence. Fragments of the HPV68 and HPV70
XX DNA may be useful as primers or probes in the detection of HPV68 or HPV70
XX infections. Proteins produced by the HPV DNA's can be used to immunise
XX animals against HPV68 or HPV70 infections
XX
XX Sequence 3283 BP; 945 A; 640 C; 699 G; 999 T; 0 U; 0 Other;
XX
Query Match 75.0%; Score 24; DB 2; Length 3283;
Best Local Similarity 84.4%; Pred. No. 1.2;
Matches 27; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 1 TTATTAATAAGTGCTGCGGTCGCAAAACC 32
```

Db 2754 TTATCATTAAGTGCATGATGTTGCTGAAACC 2785

|||||

RESULT 30

AAQ25937

ID AAQ25937 standard; DNA; 7833 BP.

XX

AC AAQ25937;

XX

DT 24-OCT-2003 (revised)

DT 25-MAR-2003 (revised)

DT 18-JAN-1993 (first entry)

XX

DE Human papilloma virus HPV39 genome.

XX

KW HPV39 infection; cancer; glucocorticoid response element; GRE; enhancer;

KW es.

XX

OS Human papillomavirus; 39.

XX

XX Key Location/Qualifiers

FT repeat_unit 43..53

FT /tag= f

FT /function= "enhancer"

FT /note= "papillomavirus-specific palindrome"

FT 58..70

FT /tag= g

FT /function= "enhancer"

FT /note= "papillomavirus-specific palindrome"

FT complement(406..411)

FT /tag= c

FT complement(424..427)

FT /tag= d

FT complement(776..2050)

FT /tag= a

FT complement(1861..1871)

FT /tag= e

FT /standard_name= "splice_acceptor_site"

FT /note= "potential"

FT /tag= b

FT complement(3875..4204)

FT /tag= h

FT /function= "enhancer"

FT /note= "papillomavirus-specific palindrome"

XX

XX MO9211369-A1.

XX

XX 09-JUL-1992.

XX

XX 20-DEC-1991; 91WO-FR001053.

XX

XX 20-DEC-1990; 90PR-00016044.

XX

PA (INSP) INST PASTEUR.

PA (INRM) INSERM INST NAT SANTE & RECH MED.

XX

PI Orth G, Volpers C, Strecek RE;

XX

DR WPI; 1992-250090/30.

XX

PT DNA sequences of the HPV39 Papilloma: virus genome - used as a probe for

PT in-vitro diagnosis of Papilloma: virus infections in cervical, vulvar and

PT penile cancer.

XX

PS Claim 1; Fig 1; 33pp; French.

XX

CC DNA was isolated from a previously identified HPV39 clone (Virology 161,

CC 374-384, 1987), subcloned and sequenced. The HPV39 genome was found to

CC comprise 7833bp and have a G/C content of 40%. The genome comprises open

CC reading frames conserved among all the known HPV's; the ORFs are

CC postulated to code for the early proteins and capsid components of the

CC virus. In addition, two ORFs have been found on the complementary strand

CC of the genome (see features table). The non-coding regions contain three

CC complete copies of an HPV enhancer as well as two degenerate versions of

CC the consensus palindrome (location not given). Other putative features

CC which are described in the specification as being present within the

CC regulatory region but whose positions are not given include: 4 nuclear

CC factor-1 binding sites, 2 activation protein-1 binding sites, a

CC papillomavirus enhancer associated factor binding site and a

CC glucocorticoid response element. The HPV39 sequence has most homology

CC with HPV18. The HPV39 is thought to belong to a sub-group of genital HPV's

CC (with HPV18, HPV45 and a new type cloned from carcinoma ME180 cell line)

CC representing a potential oncogene. (Updated on 25-MAR-2003 to correct PN

CC field.) (Updated on 24-OCT-2003 to standardise OS field)

XX

SQ Sequence 7833 BP; 2423 A; 1486 C; 1662 G; 2262 T; 0 U; 0 Other;

Query Match 75.0%; Score 24; DB 2; Length 7833;

Best Local Similarity 84.4%; Pred. No. 1.5;

Matches 27; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 TTATTAATTAAGTGCATGATGTTGCTGAAACC 32

Db 407 TTATTAATTAAGTGCATGATGTTGCTGAAACC 438

|||||

RESULT 31

AD245655

ID AD245655 standard; DNA; 7905 BP.

XX

AC AD245655;

XX

DT 30-JUN-2005 (first entry)

XX

DE Human papillomavirus type 70 full length genomic clone - SEQ ID 9.

XX

KW tumor marker; genome; diagnosis; cancer; uterine cervix tumor;

KW rectal tumor; colon tumor; de.

XX

OS Human papillomavirus type 70.

XX

PN WO200503333-A2.

XX

PD 14-APR-2005.

XX

PF 04-OCT-2004; 2004MO-DK000670.

XX

XX 07-OCT-2003; 2003DK-00001474.

PR 07-OCT-2003; 2003US-0509205P.

PR 13-FEB-2004; 2004US-0543925P.

XX

PA (DAKO-) DAKOCYTOMATION DENMARK AS.

XX

PI Chau MF, Bisgaard-Franzen K, Lin J, Raemussen OF, Wang Z, Lusk J;

PI Lindberg M, Yoast S;

XX

DR WPI; 2005-285441/29.

XX

PT New composition having a nucleic acid molecule identical to genomic clone

PT of human Papilloma virus 16, 18 or 51, useful in diagnosing cancer or the

PT risk of developing cancer, in particular cervical, anal, colon and HPV-

PT related cancer.

XX

PS Claim 2; SEQ ID NO 9; 16pp; English.

XX

CC The invention comprises a composition for the detection of cancer

CC markers, the composition contains a nucleic acid molecule which is

CC substantially identical to a full length genomic clone of a human

CC papilloma virus (type 11, 16, 18, 51, 56, 58, 66, 70, or 73). The

CC composition of the invention is useful for diagnosing cancer or the risk

CC of developing cancer, in particular cervical, anal, colon and HPV-related

CC cancer. The present DNA sequence represents a human papilloma virus full

CC length genomic clone of the invention.

PT exposed to the toxin and comparing these to gene expression in unexposed
 PT tissues or cells.
 XX
 PS Claim 1; SEQ ID NO 520; 239pp; English.
 XX
 CC The invention relates to methods for predicting toxic effects of
 CC compounds or the progression of these toxic effects by determining the
 CC global changes in gene expression in tissues or cells exposed to the
 CC toxin and comparing these to gene expression in unexposed tissues or
 CC cells. Also included are methods of predicting at least one toxic effect
 CC of a compound or progression of a toxic effect, preferably the
 CC hepatotoxicity of a compound, comprising detecting the level of
 CC expression in a tissue or cell sample exposed to the compound of two or
 CC more genes listed in the specification, where differential expression of
 CC the genes is indicative of at least one toxic effect or progression. The
 CC method can also be used to identify an agent which modulates the toxic
 CC response and predict cellular pathways that a compound modulates in a
 CC cell. The methods utilize a set of at least two probes (on a solid
 CC support in kit form), where each of the probes comprises a sequence that
 CC specifically hybridises to a gene listed in the specification, a computer
 CC system comprising a database containing information identifying the
 CC expression level in a tissue or cell sample exposed to a hepatotoxin of a
 CC set of genes comprising at least two genes listed in the specification,
 CC and a user interface to view the information used to present information,
 CC identifying the expression level in a tissue or cell of at least one gene
 CC listed in the specification. The method is useful for elucidating global
 CC changes in gene expression and for identifying toxicity markers in
 CC tissues or cell exposed to a known toxin. The genes may be used as
 CC toxicity markers in drug screening and toxicity assays. The genes and
 CC gene expression information may be used as diagnostic markers for the
 CC prediction or identification of the physiological state of tissue or cell
 CC sample that has been exposed to a compound or agent. Hepatotoxicity is
 CC characterised by centrilobular necrosis and steatosis. The present
 CC sequence is an expressed sequence tag (EST) or cDNA derived from a gene
 CC which is differentially expressed in response to a hepatotoxic agent
 CC
 SQ Sequence 630 BP; 191 A; 141 C; 171 G; 127 T; 0 U; 0 Other;
 XX
 QY Query Match 74.4%; Score 23.8; DB 6; Length 630;
 Best Local Similarity 92.6%; Pred. No. 1.1;
 Matches 25; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 DB 6 AATGAGTGCTGCGTGCCGAAACC 32
 73 AAAAAGTTCTCTGCGTGCCGAAACC 99
 XX
 RESULT 34
 ADV39914
 ID ADV39914 standard; cDNA; 630 BP.
 XX
 AC ADV39914;
 XX
 DT 10-FEB-2005 (first entry)
 XX
 DE Rat cardiotoxicity marker EST, SEQ ID NO:780.
 XX
 KW Toxicology screening; drug screening; cardiotoxicity; drug-induced;
 KW expression profile; gene expression; myocarditis; heart arrhythmia;
 KW tachycardia; myocardial ischemia; angina; hypertension; hypotension;
 KW dyspnea; cardiogenic shock; cardiovascular disease; EST;
 KW expressed sequence tag; ss.
 XX
 OS Rattus norvegicus.
 XX
 PN WO2004063334-A2.
 XX
 PD 29-JUL-2004.
 XX
 PF 08-JAN-2004; 2004WO-US000240.
 XX
 PR 08-JAN-2003; 2003US-00338044.
 XX

PA (GENE-) GENE LOGIC INC.
 XX
 PI Mendrick DL, Porter MW, Johnson KR, Higgs B, Castle A;
 PI Elashoff M;
 XX
 DR WPI; 2004-561893/54.
 XX
 DR GENBANK; A1009096.
 XX
 PT Predicting a toxic effect of a compound by obtaining a gene expression
 PT profile of a tissue or cell sample exposed to the compound and comparing
 PT the gene expression profile to a database comprising all of the data
 PT given in the specification.
 XX
 PS Example 1; SEQ ID NO 780; 520pp; English.
 XX
 CC The invention relates to a method of predicting at least one toxic effect
 CC of a compound by obtaining a gene expression profile of a tissue or cell
 CC sample exposed to the compound and comparing the gene expression profile
 CC to a database of toxicity prediction genes which are differentially
 CC expressed on exposure to known toxins. The tissue or cell sample is
 CC preferably derived from heart tissue, and the predicted toxic effect is
 CC especially a cardiotoxic effect such as myocarditis, arrhythmias,
 CC tachycardia, myocardial ischemia, angina, hypertension, hypotension,
 CC dyspnea, cardiogenic shock or other specific heart pathologies. The
 CC invention is based on the elucidation of global changes in gene
 CC expression in tissues or cells exposed to known toxins, particularly
 CC cardiotoxins, and the identification of individual genes (toxicity
 CC markers) that are differentially expressed on toxin exposure. The
 CC invention also relates to methods for predicting the progression of a
 CC compound's toxic effects, for predicting the cardiotoxicity of a
 CC compound, for identifying an agent that modulates the onset or
 CC progression of a toxic response, and for predicting which cellular
 CC pathways a particular compound will modulate. The invention further
 CC relates to sets of at least two hybridization probes specific for
 CC toxicity prediction genes, and solid supports and kits comprising them.
 CC The method of the invention is useful in toxicology screening for
 CC predicting the toxic effects (especially cardiotoxic effects) of
 CC compounds such as pharmaceutical agents or environmental pollutants.
 CC Sequences ADV39135-ADV41830 represent cardiotoxicity marker
 CC polynucleotides of rat origin whose expression is altered on exposure to
 CC at least one cardiotoxin. Note: The sequence data for this patent did not
 CC form part of the printed specification, but was obtained in electronic
 CC format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
 CC
 SQ Sequence 630 BP; 191 A; 141 C; 171 G; 127 T; 0 U; 0 Other;
 XX
 QY Query Match 74.4%; Score 23.8; DB 13; Length 630;
 Best Local Similarity 92.6%; Pred. No. 1.1;
 Matches 25; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 DB 6 AATGAGTGCTGCGTGCCGAAACC 32
 73 AAAAAGTTCTCTGCGTGCCGAAACC 99
 XX
 RESULT 35
 AAS31153
 ID AAS31153 standard; cDNA; 885 BP.
 XX
 AC AAS31153;
 XX
 DT 04-DEC-2001 (first entry)
 XX
 DE Human diagnostic and therapeutic polynucleotide (DTRP) #168.
 XX
 KW Human; receptor; diagnostic; therapeutic; gene therapy; vaccine;
 KW cell proliferative disorder; Crohn's disease; lymphoma; leukaemia;
 KW acquired immune deficiency syndrome; AIDS; autoimmune disorder;
 KW respiratory disorder; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200162927-A2.
 XX

XX 30-AUG-2001.
 PD 21-FEB-2001; 2001MO-US006059.
 XX
 XX 24-FEB-2000; 2000US-0184693P.
 PR 24-FEB-2000; 2000US-0184697P.
 PR 24-FEB-2000; 2000US-0184698P.
 PR 24-FEB-2000; 2000US-0184768P.
 PR 24-FEB-2000; 2000US-0184769P.
 PR 24-FEB-2000; 2000US-0184770P.
 PR 24-FEB-2000; 2000US-0184771P.
 PR 24-FEB-2000; 2000US-0184772P.
 PR 24-FEB-2000; 2000US-0184773P.
 PR 24-FEB-2000; 2000US-0184774P.
 PR 24-FEB-2000; 2000US-0184776P.
 PR 24-FEB-2000; 2000US-0184777P.
 PR 24-FEB-2000; 2000US-0184779P.
 PR 24-FEB-2000; 2000US-0184813P.
 PR 24-FEB-2000; 2000US-0184817P.
 PR 24-FEB-2000; 2000US-0185213P.
 PR 24-FEB-2000; 2000US-0185216P.
 PR 12-MAY-2000; 2000US-0203785P.
 PR 15-MAY-2000; 2000US-0204226P.
 PR 16-MAY-2000; 2000US-0204525P.
 PR 16-MAY-2000; 2000US-0204821P.
 PR 16-MAY-2000; 2000US-0204908P.
 PR 16-MAY-2000; 2000US-0205232P.
 PR 17-MAY-2000; 2000US-0204815P.
 PR 17-MAY-2000; 2000US-0204863P.
 PR 17-MAY-2000; 2000US-0205221P.
 PR 17-MAY-2000; 2000US-0205285P.
 PR 17-MAY-2000; 2000US-0205286P.
 PR 17-MAY-2000; 2000US-0205287P.
 PR 17-MAY-2000; 2000US-0205323P.
 PR 17-MAY-2000; 2000US-0205324P.
 XX
 PA (INCY-) INCYTE GENOMICS INC.
 XX
 PI Panzer SR, Spiro PA, Banville SC, Shah P, Chalup MS, Chang SC;
 PI Chen A, D'sa SA, Amshay S, Dahl CR, Dam TC, Daniels SE, Dutoir GB;
 PI Flores V, Fong WT, Greenwalt LB, Hillman JL, Jones AL, Liu TF;
 PI Roseberry AM, Rosen BH, Russo FD, Stockdreher TK, Dafio A;
 PI Wright RJ, Yap PR, Yu JY, Bradley DL, Bratcher SR, Chen W;
 PI Cohen HU, Hodgson DM, Lincoln SE, Jackson S;
 XX
 DR WPI, 2001-502867/55.
 DR P-PSDB; AAU19582.
 XX
 PT Polynucleotides encoding diagnostic and therapeutic proteins, e.g.
 PT enzymes, hormones and receptors, useful in diagnostics and therapeutics.
 XX
 PS Claim 1, Page 378; 522pp; English.
 XX
 CC The invention relates to polynucleotides (I) encoding diagnostic and
 CC therapeutic (DITHP) polypeptides (II), which include e.g. enzymes, and
 CC proteins involved in growth and development and receptors. (I) and (II)
 CC may be used in the prevention, diagnosis and treatment of diseases
 CC associated with inappropriate DITHP expression. For example, (I) and (II)
 CC may be used to treat disorders associated with decreased polypeptide
 CC expression by rectifying mutations or deletions in a patient's genome,
 CC that affect the activity of the DITHPs, by expressing inactive proteins
 CC or supplementing the patient's own production of them. (I) and (II) may
 CC be used to treat diseases, for example, cell proliferative disorder,
 CC Crohn's disease, acquired immune deficiency syndrome (AIDS), lymphoma,
 CC leukemia, autoimmune disorders, and respiratory disorders. Additionally,
 CC (I) may be used to produce the DITHPs, by inserting the nucleic acids
 CC into a host cell and culturing the cell to express the protein. (I) and
 CC its complementary sequences may also be used as DNA probes in diagnostic
 CC assays to detect and quantitate the presence of similar nucleic acids in
 CC samples, and therefore which patients may be in need of restorative
 CC therapy. (II) may also be used as antigens in the production of

CC antibodies against DITHPs and in assays to identify modulators of DITHP
 CC expression and activity. The anti-DITHP antibodies and antagonists may
 CC also be used to down regulate expression and activity. The anti-DITHP
 CC antibodies may also be used as diagnostic agents for detecting the
 CC presence of DITHPs in samples (e.g. by enzyme linked immunosorbent assay
 CC (ELISA)). AAS30986-AAS31196 represent human diagnostic and therapeutic
 CC (DITHP) polynucleotides of the invention
 XX
 SQ Sequence 885 BP; 284 A; 181 C; 229 G; 187 T; 0 U; 4 Other;
 Query Match 74.4%; Score 23.8; DB 4; Length 885;
 Best Local Similarity 92.6%; Pred. No. 1.2;
 Matches 25; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Oy 6 AATAGAGTGCCTGCGGTGCAGAAACC 32
 Db 69 AAAAGATTCTCTCGGTGCCAGAAACC 95
 RESULT 36
 AAT71196/c
 ID AAT71196 standard; DNA; 40 BP.
 XX
 AC AAT71196;
 XX
 DT 25-MAR-2003 (revised)
 DT 20-AUG-1997 (first entry)
 XX
 DE HPV type 18 exon-specific probe, binds to spliced B6 region.
 XX
 KW Human papillomavirus; HPV; cervical cancer; cervix; neoplasia; tumour;
 KM probe; detection; diagnosis; hybridisation; ss.
 XX
 OS Synthetic.
 XX
 EN US5580970-A.
 XX
 PD 03-DEC-1996.
 XX
 PF 04-MAR-1994; 94US-00207226.
 XX
 PR 01-DEC-1989; 89US-00444526.
 PR 05-DEC-1990; 90US-00622742.
 XX
 PA (STAD) AMOCO CORP.
 XX
 PI Parodos K, Rigby S, Lane DJ, Hendricks DA;
 XX
 DR WPI, 1997-033604/03.
 XX
 PT Nucleic acid probes specific for human papilloma virus - for prognosis of
 PT cervical neoplasias and cancers.
 XX
 PS Example 1; Fig 4; 18pp; English.
 XX
 CC AAT71194-T71199 are probes used for the detection of spliced exons of
 CC human papillomavirus (HPV) types 16, 18, 31, 33 and 35. The probes
 CC specifically hybridise to B6 region exons at the splice junction of the
 CC spliced HPV genome. The probes are used for accurate and reliable
 CC prognosis of HPV-induced cancers, in particular cervical cancer. (Updated
 CC on 25-MAR-2003 to correct PF field.)
 XX
 SQ Sequence 40 BP; 8 A; 12 C; 9 G; 11 T; 0 U; 0 Other;
 Query Match 72.5%; Score 23.2; DB 2; Length 40;
 Best Local Similarity 89.3%; Pred. No. 1.2;
 Matches 25; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 Oy 4 TTATATAGGTGCTGCGGTGCCAGAAAC 31
 Db 28 TTACAGAGTGTCTGCGGTGCCAGAAAC 1

```

RESULT 37
AAQ44290/c
ID AAQ44290 standard; DNA; 30 BP.
XX
AC AAQ44290;
XX
DT 25-MAR-2003 (revised)
DT 02-DEC-1993 (first entry)
XX
DE Sequence of primer 2PV18 5' (C6, NH2) for human papilloma virus (HPV)
DE directed to the E6 region (19).
XX
KM Primer; exponential amplification; PCR; human papilloma virus; HPV18; ss.
XX
OS Synthetic.
XX
PN WO9312245-A1.
XX
PD 24-JUN-1993.
XX
PF 07-DEC-1992; 92MO-US010524.
XX
PR 11-DEC-1991; 91US-00804951.
XX
PA (IGEN-) IGEN INC.
XX
PI Kenten JH, Link JR;
XX
DR WPI; 1993-214186/26.
XX
PT Amplifying nucleic acid sequences of interest - using single primer to
PT provide exponential amplification in polymerase reaction.
XX
PS Example; Page 32; 54pp; English.
XX
CC The invention is in a process for exponentially amplifying a nucleic acid
CC sequence by a single unpaired primer. It includes a method for modifying
CC a putative primer to improve its ability to prime a single primer
CC exponential amplification process. Oligo amino modifications to the 5'
CC end occurred at the last coupling step. The resulting 5' modified oligos
CC all contained a six carbon spacer arm to the amino gp., designated (C6,
CC NH2). Oligo Ru(II) modifications to the 5' end occurred at the last
CC coupling step and the oligo is designated Ru(II) (Updated on 25-MAR-2003
CC to correct RN field.)
XX
SQ Sequence 30 BP; 10 A; 7 C; 4 G; 9 T; 0 U; 0 Other;
XX
Query Match 71.9%; Score 23; DB 2; Length 30;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 TTATTATAAGTGCTCGCGTG 23
Db 23 TTATTATAAGTGCTCGCGTG 1

```

```

PD 05-OCT-1999.
XX
XX 24-NOV-1993; 93US-00160063.
XX
XX 03-NOV-1988; 88US-00266882.
PR 18-JUN-1990; 90US-00539389.
PR 06-FEB-1991; 91US-00652427.
PR 03-FEB-1992; 92US-00827270.
PR 12-JUL-1993; 93US-00090467.
PR 24-NOV-1993; 93US-00158193.
XX
XX (IGEN-) IGEN INT INC.
XX
PI Shah HP, Kenten JH, Goodman JE, Blackburn GF, Maesey RJ;
PI Lowke GE, Leland JK;
XX
XX WPI; 1999-579609/49.
XX
PT Binding assay based on detection of electrochemiluminescence.
XX
XX Example 20; Col 30; 43pp; English.
XX
PS The invention provides a binding assay based on measuring
XX CC electrochemiluminescence from a luminescent component concentrated in an
XX CC assay composition which is collected at an electrode for detection. The
XX CC binding assay is used to measure luminescence emitted by one or more
XX CC labeled components of the assay system. The assay can be used to measure
XX CC e.g. antigen-antibody, ligand-receptor, RNA and DNA interactions. The
XX CC sample may be derived from cells, water, feces, saliva, tissue, hair,
XX CC sweat or urine and the analyte of interest may be e.g. a whole cell,
XX CC surface antigen, subcellular particle, virus, antibody, antigen, hapten,
XX CC polysaccharide, hormone, nucleic acid, vitamin, amino acid, sugar,
XX CC recombinant protein, pharmacological agent, steroid or inorganic
XX CC molecule. The assay has improved sensitivity, faster assay time, greater
XX CC specificity, lower detection limits and greater precision than previous
XX CC methods. Sequences AAZ07818-823 represent PCR oligos directed to the HPV
XX CC E6 region, used during the course of the invention
XX
SQ Sequence 30 BP; 10 A; 7 C; 4 G; 9 T; 0 U; 0 Other;
XX
Query Match 71.9%; Score 23; DB 2; Length 30;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 TTATTATAAGTGCTCGCGTG 23
Db 23 TTATTATAAGTGCTCGCGTG 1

```

```

RESULT 39
ADZ64772/c
ID ADZ64772 standard; DNA; 30 BP.
XX
AC ADZ64772;
XX
DT 30-JUN-2005 (first entry)
XX
DE HPV 18 E6 region specific PCR primer, 2PV18.
XX
XX Analyte detection; luminescence; PCR; primer; ss.
XX
XX Human papillomavirus type 18.
XX
FH Key Location/Qualifiers
FT modified_base 1
FT /*tag= a
FT /mod_base= OTHER
FT /note= "linked to a six carbon spacer arm designated (C6,
FT NH2)"
FT modified_base 30
FT /*tag= b
FT /mod_base= OTHER
FT /note= "linked to a three carbon spacer"
FT

```

```

XX  US6881589-B1.
XX
XX  19-APR-2005.
XX
XX  05-JUN-1995; 95US-00465443.
XX
XX  30-APR-1987; 87WO-US000987.
XX
XX  03-NOV-1988; 88US-00266882.
XX
XX  21-JUL-1989; 89US-00369560.
XX
XX  18-JUN-1990; 90US-00539389.
XX
XX  06-FEB-1991; 91US-00652427.
XX
XX  24-NOV-1993; 93US-00158193.
XX
XX  30-NOV-1994; 94US-00346832.
XX
XX  30-APR-1996; 96US-00858354.
XX
XX  (BIOV-) BIOVERIS CORP.
XX
XX  Ireland UK, Shah HP, Kenten JH, Goodman JE, Lowke GE, Namba Y,
XX  Blackburn GF, Massey RJ;
XX
XX  WPI; 2005-312301/32.
XX
XX  Electroluminescent composition useful in assay for analyte e.g.,
XX  virus, comprises electroluminescent label, first and second binding
XX  reagents having binding partners of analyte, collectable magnetic
XX  particles and electrolyte.
XX
XX  Example 20; Col 29; 48pp; English.
XX
XX  The present invention relates to an electroluminescent (EL)
XX  composition suitable for use in electroluminescent assay for analyte
XX  of interest which comprises electroluminescent label binding
XX  reagents, collectable magnetic particles and an electrolyte. The present
XX  sequence is the human papillomavirus (HPV) E6 region specific PCR primer.
XX  This sequence was used in a binding assay for an analyte of interest
XX  present in a sample based upon electroluminescence.
XX
XX  Sequence 30 BP; 10 A; 7 C; 4 G; 9 T; 0 U; 0 Other;
XX
XX  Query Match 71.9%; Score 23; DB 14; Length 30;
XX  Best Local Similarity 100.0%; Pred. No. 1.4;
XX  Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX  1 TTTATTAATAGTGCTGCGGTG 23
XX  23 TTATTAATAGTGCTGCGGTG 1
XX
XX  RESULT 40
XX  ABV07997
XX  ID ABV07997 standard; cDNA; 167 BP.
XX
XX  ABV07997;
XX
XX  13-SBP-2002 (first entry)
XX
XX  Human prostate expression marker cDNA 7988.
XX
XX  Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
XX  pharmacogenomic marker; gene; ss.
XX
XX  Homo sapiens.
XX
XX  WO200160860-A2.
XX
XX  23-AUG-2001.
XX
XX  20-FEB-2001; 2001WO-US005171.
XX
XX  17-FEB-2000; 2000US-0183319P.
XX
XX  16-MAR-2000; 2000US-0189862P.
XX
XX  25-MAY-2000; 2000US-0207454P.

```

```

PR  09-JUN-2000; 2000US-0211314P.
PR  18-JUL-2000; 2000US-0219007P.
PR  13-DEC-2000; 2000US-025281P.
XX
XX  (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
XX  Schlegel R, Endege WO, Monahan JR;
XX
XX  WPI; 2001-662795/76.
XX
XX  Novel isolated nucleic acid molecule associated with cancerous state of
XX  prostate cells and correlating with presence of prostate cancer, useful
XX  for detecting presence of prostate cancer, stage of prostate cancer.
XX
XX  Claim 1; Page 1279; 11750pp; English.
XX
XX  The invention relates to an isolated nucleic acid molecule (I) comprising
XX  a nucleotide sequence given in Tables 1-9 (ABV0010-ABV62213) of the
XX  specification or its complement. (I) is useful for: (a) assessing whether
XX  a patient is afflicted with prostate cancer; (b) monitoring the
XX  progression of prostate cancer in a patient; (c) assessing the efficacy
XX  of a test compound to inhibit prostate cancer in a patient; (d) assessing
XX  the efficacy of a therapy for inhibiting prostate cancer in a patient;
XX  (e) selecting a composition for inhibiting prostate cancer in a patient;
XX  (f) assessing the prostate cell carcinogenic potential of a compound; (g)
XX  determining whether prostate cancer has metastasized in a patient; (h)
XX  assessing the aggressiveness or indolence of prostate cancer in a patient
XX  ; (i) is also useful as a pharmacodynamic or pharmacogenomic marker
XX
XX  Sequence 167 BP; 48 A; 33 C; 48 G; 37 T; 0 U; 1 Other;
XX
XX  Query Match 70.6%; Score 22.6; DB 5; Length 167;
XX  Best Local Similarity 83.3%; Pred. No. 3;
XX  Matches 25; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
XX
XX  3 ATTAATAGTGCTGCGGTGCCGAAACC 32
XX  88 AANAAGAGGTCTCTGCTGCCAGAAACC 117
XX
XX  RESULT 41
XX  ACH89313/C
XX  ID ACH89313 standard; DNA; 159 BP.
XX
XX  ACH89313;
XX
XX  29-JUL-2004 (first entry)
XX
XX  Human genome derived single exon probe #22508.
XX
XX  Human; probe; ss; gene expression; single exon probe; microarray;
XX  alternative splicing event; genomic alteration.
XX
XX  Homo sapiens.
XX
XX  US2003194704-A1.
XX
XX  16-OCT-2003.
XX
XX  03-APR-2002; 2002US-00029386.
XX
XX  03-APR-2002; 2002US-00029386.
XX
XX  (PENN/) PENN S G.
XX  (RANK/) RANK D R.
XX  (HANZ/) HANZEL D K.
XX
XX  Penn SG, Rank DR, Hanzel DK;
XX
XX  WPI; 2004-119264/12.
XX
XX  New human genome-derived single exon nucleic acid probes useful for human
XX  gene expression analysis, for identifying or characterizing alternative

```

PT splicing events, for assessing genomic alterations or as tools for PT surveying tissues.

PS Claim 1; SEQ ID NO 22508; 80pp; English

The invention relates to a nucleic acid probe for measuring human gene expression, comprising any of the 27,400 fully defined or nucleotide sequences in the specification, or their complements or fragments, and encoding at least 8 amino acids of any of the 688 amino acid sequences fully defined in the specification. The probe is a single exon probe that hybridises under high stringency conditions to a nucleic acid molecule expressed in human cells or tissues. Also included are a spatially-addressable set of single exon nucleic acid probes for measuring human gene expression (comprising a plurality of single exon nucleic acid probes cited above, where each of the plurality of probes is separately and addressably isolatable or amplifiable from the plurality), a single exon microarray for measuring human gene expression, a method of measuring human gene expression, a vector comprising the single exon probe cited above, an ORF-encoded peptide comprising at least 8 contiguous amino acids of any of the above-mentioned amino acid sequences (optionally with conservative amino acid substitutions), an isolated antibody that binds specifically to a peptide cited above, methods of selling and/or licensing single exon probes or microarrays to a customer desiring to measure gene expression, a method of providing human gene expression data by subscription, and a computer-readable storage medium which contains a database having a plurality of records (each record including data on the expression of a single exon probe cited above. The probe, methods and apparatus are useful in gene expression analysis. The probes may be used as tools for surveying tissues to detect the presence of expressed messages that contain their specific exon, or in constructing genome-derived single exon microarrays. In addition, the probes are used in identifying and characterising alternative splicing events, in detecting and characterising gross alterations in the genomic locus that includes their exon, in assessing smaller genomic alterations, in priming the synthesis of nucleic acids, or in expressing the ORF-encoded peptide. The present sequence is a human single exon probe of the invention. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from USPTO at seqdata.uspto.gov/sequence.html?DocID=20030194704

Seq	Sequence	159 BP	30 A	40 C	26 G	63 T	0 U	0 other
Qy	Query Match	69.4%						
	Similarity	88.9%						
	Best Local	Pred. No. 2.2	DB 12					
	Matches	24	Conservative	0	Mismatches	3	Indels	0
								Gaps 0
Ob	6 AATAAGGTGCTGCGCGTCCAGAAACC	32						
	127 AAGAAGTTCTCTGCTGTCAGAAACC	101						

```

RESULT 42
AAS78228
ID AAS78228 standard; cDNA; 297 BP.

```

AC	AA578228;
XX	
DT	13-FEB-2002 (first entry)
XX	
DE	DNA encoding novel human diagnostic protein #14032

KM Human; chromosome mapping; gene mapping; gene therapy; forensic; food supplement; medical imaging; diagnostic; genetic disorder; ss

Os Homo sapiens.

PN WO200175067-A2.

PD 11-OCT-2001

30-MAR-2001; 2001WO-US008631.

XX

PR 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.

PA (HYSE-) HYSEQ INC.

PI Dymanac RT, Liu C, Tang YT;

DR WPI; 2001-639362/73.

DR P-PSDB; ABG14041.

PT New isolated polynucleotide and encoded polypeptides, useful in PT diagnostics, forensics, gene mapping, identification of mutations PT responsible for genetic disorders or other traits and to assess PT biodiversity.

PS Claim 1; SEQ ID NO 14032; 103pp; English

The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. AB664197-AB594664 represent novel human diagnostic coding sequences of the invention. Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at [ftp.wipo.int/publ/published/pct_sequences](http://wipo.int/publ/published/pct_sequences)

Query Match	69.4%	Score 22.2	DB 5	Length 297
Best Local Similarity	88.9%	Pred. No. 5.2		
Matches 24	Conservative 0	Mismatches 3	Indels 0	Gaps 0

Oy 6 AATAAGTGCCTGCGGTGCCAGAAACC 32
 |||||
 Db 59 AAGAAAGTTCCTGCTGTGCCAGAAACC 85

RESULT 43
ABX40477
ID ABX40477 standard; cDNA; 421 BP.

AC ABX40477;

DT 20-FEB-2003 (first entry)

DE Bovine EST associated with lactation/muscle/fat deposition #5642.

KW Bovine; 88; EST; expressed sequence tag; lactation; LMPD;

gene analysis; cattle breeding.

OS Bos Taurus.

US2002137139-A1

26-SEP-2002

AA 24-SEP-2001
PF

XX 12-JAN-1999. 0911S-0115707P
 DP

PK 12-JAN-1999; 9905-0113/0/E

```

PR 11-JAN-2000; 2000US-00480902.
XX
XX (BYAT/) BYATT J C.
PA (MATH/) MATHIALAGAN N.
PA (TAON/) TAO N.
PA (WARR/) WARREN W C.
XX
XX Byatt JC, Mathialagan N, Tao N, Warren WC;
XX WPI; 2003-110599/10.
XX
XX New nucleic acid associated with lactation, and muscle and fat
PT deposition, useful for genome mapping, gene identification and analysis,
PT cattle breeding, or for genetically improving cattle.
XX
XX Claim 2; SEQ ID NO 5642; 245bp; English.
XX
XX The invention relates to a purified nucleic acid molecule associated with
CC lactation or muscle and fat deposition (designated LMPD), derived from
CC cattle, and the LMPD nucleic acid can specifically hybridize to a second
CC nucleic acid molecule comprising any of 1512 nucleotide sequences,
CC appearing as ABX34836-ABX49947, or complements of them. Also included are
CC (1) a transformed cell having a nucleic acid comprising an LMPD nucleic
CC acid linked to a promoter and a 3', non-translated sequence that
CC functions in the cell to cause termination of transcription and addition
CC of polyadenylated ribonucleotides to a 3' end of the mRNA molecule; and
CC (2) determining a level or pattern of a molecule in a bovine cell or
CC tissue comprising: (a) incubating a marker nucleic acid (comprising any
CC of the 1512 nucleic acid sequences or its complement or fragment) with a
CC complementary nucleic acid molecule obtained from the bovine cell or
CC tissue, where hybridization between the marker nucleic acid and the
CC complementary nucleic acid permits the detection of the molecule; and (b)
CC detecting the level or pattern of the complementary nucleic acid, where
CC the detection of the complementary nucleic acid is predictive of the
CC level or pattern of the molecule. The LMPD nucleic acid is used for
CC determining a level or pattern of a molecule in a bovine cell or tissue.
CC It is useful for genome mapping, gene identification and analysis, cattle
CC breeding, preparation of constructs for use in cattle gene expression, or
CC for genetically improving cattle. The present sequence is one of the
CC 1512 bovine LMPD BST (expressed sequence tag) nucleic acids. Note: The
CC present sequence was not shown in the specification but was obtained in
CC electronic format from the USPTO web site:
CC seqdata.uspto.gov/sequence.html?DocID=20020137139
XX
XX Sequence 421 BP; 142 A; 85 C; 111 G; 83 T; 0 U; 0 Other;
SQ
XX
XX Query Match 69.4%; Score 22.2; DB 8; Length 421;
XX Best Local Similarity 88.9%; Pred. No. 5.6;
XX Matches 24; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 6 AATAAGTGCTCGGTCGCGAACC 32
DB 30 AAGAAGTCTCTGCTGTGCCGAACC 56

```

```

XX
XX 30-JAN-2001; 2001WO-US000670.
XX
XX 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-488901/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human cervical epithelial cells.
XX
XX Claim 25; SEQ ID NO 62; 487bp; English.
XX
XX The present invention relates to human single exon nucleic acid probes
CC (SENPs). The present sequence is one such probe. The SENPs are derived
CC from human HeLa cells. The SENPs can be used to produce a single exon
CC microarray, which can be used for measuring human gene expression in a
CC sample derived from human cervical epithelial cells. By measuring gene
CC expression, the probes are therefore useful in grading and/or staging of
CC diseases of the cervix, notably cervical cancer. Note: The sequence data
CC for this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 460 BP; 163 A; 87 C; 112 G; 98 T; 0 U; 0 Other;
SQ
XX
XX Query Match 69.4%; Score 22.2; DB 4; Length 460;
XX Best Local Similarity 88.9%; Pred. No. 5.7;
XX Matches 24; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 6 AATAAGTGCTCGGTCGCGAACC 32
DB 38 AAGAAGTCTCTGCTGTGCCGAACC 64

```

```

RESULT 45
ABAS1759
XX ID ABAS1759 standard; DNA; 460 BP.
XX AC ABAS1759;
XX
XX 01-FEB-2002 (first entry)
DT 01-FEB-2002 (first entry)
XX
XX Human foetal liver single exon nucleic acid probe #64.
DE Human foetal liver single exon nucleic acid probe; ss.
XX
XX Human, foetal liver; gene expression; single exon nucleic acid probe; ss.
XX
XX Homo sapiens.
OS
XX
XX WO200157277-A2.
PN
XX
XX 09-AUG-2001.
PD
XX
XX 30-JAN-2001; 2001WO-US000669.
PF
XX
XX 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
PA

```

XX Penn SG, Hanzel DK, Chen W, Rank DR;
PI WPI; 2001-483447/52.
XX Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human fetal liver.
XX
PS Claim 1; SEQ ID NO 64; 639pp + Sequence Listing; English.
XX
XX The invention relates to a single exon nucleic acid probe for measuring
CC human gene expression in a sample derived from human foetal liver. The
CC single exon nucleic acid probes may be used for predicting, measuring and
CC displaying gene expression in samples derived from human fetal liver. The
CC present sequence is a single exon nucleic acid probe of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 460 BP; 163 A; 87 C; 112 G; 98 T; 0 U; 0 Other;
XX
Query Match 69.4%; Score 22.2; DB 4; Length 460;
Best Local Similarity 88.9%; Pred. No. 5.7;
Matches 24; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
OY 6 AATPAGGTGCTGCGTGCAGAAACC 32
DB 38 AAGAAGTCTCTGCTGCGAGAAACC 64
RESULT 46
AA131375
ID AA131375 standard; DNA; 460 BP.
XX
AC AA131375;
XX
DT 17-OCT-2001 (first entry)
XX
DE Probe #61 used to measure gene expression in human placenta sample.
XX
KM Probe; microarray; human; placenta; antenatal diagnosis;
KM genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200157272-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000663.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488897/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human placenta.
XX
PS Claim 25; SEQ ID NO 61; 654pp; English.
XX
CC The present invention relates to single exon nucleic acid probes (SENP).
CC The present sequence is one such probe. The probes are useful for
CC producing a microarray for predicting, measuring and displaying gene

CC expression in samples derived from human placenta. The probes are useful
CC for antenatal diagnosis of human genetic disorders
XX
SQ Sequence 460 BP; 163 A; 87 C; 112 G; 98 T; 0 U; 0 Other;
XX
Query Match 69.4%; Score 22.2; DB 4; Length 460;
Best Local Similarity 88.9%; Pred. No. 5.7;
Matches 24; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
OY 6 AATPAGGTGCTGCGTGCAGAAACC 32
DB 38 AAGAAGTCTCTGCTGCGAGAAACC 64
RESULT 47
ABA21587
ID ABA21587 standard; DNA; 460 BP.
XX
AC ABA21587;
XX
DT 23-JAN-2002 (first entry)
XX
DE Probe #53 for gene expression analysis in human heart cell sample.
XX
KM Human; gene expression; heart; microarray; vascular system; probe;
KM cardiovascular disease; hypertension; cardiac arrhythmia;
KM congenital heart disease; ss.
XX
OS Homo sapiens.
XX
PN WO200157274-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000666.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488899/53.
XX
PT Single exon nucleic acid probes for analyzing gene expression in human
PT hearts.
XX
PS Claim 1; SEQ ID NO 53; 530pp; English.
XX
CC The present invention relates to single exon nucleic acid probes for
CC measuring human gene expression in a sample derived from human heart. The
CC present sequence is one such probe. The probes may be used for
CC predicting, measuring and displaying gene expression in samples derived
CC from the human heart via microarray. By measuring gene expression, the
CC probes are useful for predicting, diagnosing, grading, staging,
CC monitoring and prognosticating diseases of the human heart and vascular system
CC e.g. cardiovascular disease, hypertension, cardiac arrhythmias and
CC congenital heart disease. Note: The sequence data for this patent did not
CC form part of the printed specification, but was obtained in electronic
CC format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 460 BP; 163 A; 87 C; 112 G; 98 T; 0 U; 0 Other;
XX
Query Match 69.4%; Score 22.2; DB 4; Length 460;
Best Local Similarity 88.9%; Pred. No. 5.7;
Matches 24; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

PR 26-MAY-2000; 2000US-0207456P.
 PR 30-JUN-2000; 2000US-00608408.
 PR 03-AUG-2000; 2000US-00632366.
 PR 21-SEP-2000; 2000US-0234687P.
 PR 27-SEP-2000; 2000US-0236359P.
 PR 04-OCT-2000; 2000GB-00024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-488898/53.

PT Human genome-derived single exon nucleic acid probes useful for analyzing
 PT gene expression in human adult liver.

PS Claim 1; SEQ ID NO 63; 658bp; English.

XX The invention relates to a single exon nucleic acid probe (SENP) (I) for
 CC measuring human gene expression in a sample derived from human adult
 CC liver, comprising one of 13109 defined nucleotide sequences given in the
 CC specification (or complements/ fragments). The probe hybridises at high
 CC stringency to a nucleic acid molecule expressed in the human adult liver.
 CC (I) may be used for predicting, measuring and displaying gene expression
 CC in samples derived from human adult liver. The genes identified may be
 CC involved in genetic liver diseases such as cirrhosis.
 CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is
 CC associated with coronary heart disease. ABS25011-ABS51005 represent human
 CC liver single exon nucleic acid probes of the invention. Note: The
 CC sequence information for this patent does not appear in the printed
 CC specification but was obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 460 BP; 163 A; 87 C; 112 G; 98 T; 0 U; 0 Other;

Query Match 69.4%; Score 22.2; DB 4; Length 460;

Best Local Similarity 88.9%; Pred. No. 5.7;
 Matches 24; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 6 AATAGGTGCTGCGGTGCCAGAAACC 32

DB 38 AAGAAAGTCTGCTGCTGCCAGAAACC 64

Search completed: May 24, 2006, 05:55:19
 Job time : 382.899 secs

THIS PAGE IS BLANK

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 24, 2006, 05:47:01 ; Search time 2208.88 Seconds
(without alignments)
810.102 Million cell updates/sec

Title: US-10-601-913-121

Perfect score: 32
Sequence: 1 TTTATATATAGTGCTGCGGTGCCAGAAACC 32

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 48236798 seqs, 2795965780 residues

Total number of hits satisfying chosen parameters: 96473596

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 1000 summaries

Database :

EST.*
1: gb_est1:*
2: gb_est3:*
3: gb_est4:*
4: gb_est5:*
5: gb_est6:*
6: gb_est7:*
7: gb_est8:*
8: gb_est9:*
9: gb_est10:*
10: gb_est11:*
11: gb_gsa1:*
12: gb_gsa2:*
13: gb_gsa3:*
14: gb_gsa4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	32	100.0	316	10	DM414483 HHAGB0152
2	32	100.0	549	9	DA567416
3	32	100.0	890	7	BE543005
4	25.6	80.0	443	2	BE911801
5	24.2	75.6	219	10	DM445957
6	24.2	75.6	229	10	DM443699
7	24.2	75.6	287	10	DM424800
8	24.2	75.6	289	10	DM424288
9	24.2	75.6	294	10	DM422532
10	24.2	75.6	312	10	DM415682
11	24.2	75.6	405	2	BE912441
12	24.2	75.6	405	2	BE912537
13	24.2	75.6	407	10	DM419176
14	24.2	75.6	410	2	BE912279
15	24.2	75.6	420	7	AM675319
16	24.2	75.6	509	7	BE018031
17	24.2	75.6	742	2	BI260198
18	24.2	75.6	748	2	BI333226
19	24.2	75.6	758	2	BI222960

20	24.2	75.6	827	2	BI255152	BI255152 602977434
21	24.2	75.6	833	2	BI223045	BI223045 602943381
22	24.2	75.6	833	4	BX400902	BX400902 BX400902
23	24.2	75.6	856	7	BE544557	BE544557 601076564
24	24.2	75.6	922	7	BE544755	BE544755 601075520
25	24.2	75.6	1043	2	BI256241	BI256241 602975055
26	24.2	75.6	1114	2	BI257598	BI257598 60297170
27	24	75.0	554	1	AI747399	AI747399 u105h08.y
28	23.8	74.4	202	1	AA68665	AA68665 EST110976
29	23.8	74.4	212	1	AA686627	AA686627 EST110942
30	23.8	74.4	289	10	DM320563	DM320563 LRAGB0453
31	23.8	74.4	299	1	AA685796	AA685796 EST108189
32	23.8	74.4	323	10	DM300318	DM300318 LRAGB02621
33	23.8	74.4	341	1	AA685923	AA685923 EST108644
34	23.8	74.4	344	7	BE127031	BE127031 DEPA0780
35	23.8	74.4	346	1	AI229881	AI229881 EST226576
36	23.8	74.4	346	10	H35039	H35039 EST109541 R
37	23.8	74.4	350	10	DM380853	DM380853 LRAGB0255
38	23.8	74.4	359	1	AJ647449	AJ647449 AJ647449
39	23.8	74.4	359	1	AJ647449	BE128703
40	23.8	74.4	367	7	BE128703	CV971013
41	23.8	74.4	386	8	CV971013	DM383713 LRAGB0287
42	23.8	74.4	386	10	DM383713	DM383827 LRAGB0288
43	23.8	74.4	387	10	DM383827	DM385499 LRAGB0296
44	23.8	74.4	398	10	DM385499	DM386589 LRAGB0299
45	23.8	74.4	401	10	DM386589	AA956776 UI-R-B1-f
46	23.8	74.4	415	5	AA956776	CF273504 IGAM8 Rac
47	23.8	74.4	478	5	CF273504	AA944687 EST200186
48	23.8	74.4	490	1	AA944687	AJ647149 AJ647149
49	23.8	74.4	498	1	AJ647149	AJ650336 AJ650336
50	23.8	74.4	502	4	CB784201	CB784201 AMGNNUC:C
51	23.8	74.4	507	5	CB784201	CB784201 UI-R-GRO-
52	23.8	74.4	507	5	CB784201	CB784201 AMGNNUC:C
53	23.8	74.4	513	4	CB71861	CB71861 AMGNNUC:C
54	23.8	74.4	519	7	BE408050	BE408050 UI-R-BJ2-
55	23.8	74.4	521	4	CB615525	CB615525 AMGNNUC:N
56	23.8	74.4	574	4	CB608193	CB608193 AMGNNUC:N
57	23.8	74.4	592	5	CD372858	CD372858 UI-R-GRO-
58	23.8	74.4	605	4	CA339837	CA339837 NISC 1y08
59	23.8	74.4	607	5	CF273502	CF273502 IGA7368 R
60	23.8	74.4	619	3	BE94183	BE94183 BP494183
61	23.8	74.4	620	10	DV216610	DV216610 gx25e06.y
62	23.8	74.4	620	3	BU488468	BU488468 604124981
63	23.8	74.4	621	4	CB578714	CB578714 AMGNNUC:U
64	23.8	74.4	624	1	AM914081	AM914081 EST745398
65	23.8	74.4	630	1	AI100906	AI100906 EST703547
66	23.8	74.4	631	10	DM398288	DM398288 LRAGB0406
67	23.8	74.4	632	5	CD372845	CD372845 UI-R-GRO-
68	23.8	74.4	637	4	CB576783	CB576783 AMGNNUC:C
69	23.8	74.4	638	3	BP500480	BP500480 BP500480
70	23.8	74.4	651	5	CD372782	CD372782 UI-R-GRO-
71	23.8	74.4	654	10	DT901616	DT901616 1480859 M
72	23.8	74.4	666	4	CB557077	CB557077 AMGNNUC:N
73	23.8	74.4	668	2	BM388719	BM388719 UI-R-DZ0-
74	23.8	74.4	671	5	CF977555	CF977555 PAR1281.R
75	23.8	74.4	702	5	CK366084	CK366084 AGENCOURT
76	23.8	74.4	706	10	DT897178	DT897178 1459354 M
77	23.8	74.4	712	5	CK365811	CK365811 AGENCOURT
78	23.8	74.4	739	3	BU491662	BU491662 604129275
79	23.8	74.4	741	8	CV120816	CV120816 AGENCOURT
80	23.8	74.4	743	8	CO562425	CO562425 AGENCOURT
81	23.8	74.4	744	3	BU492327	BU492327 604129172
82	23.8	74.4	761	5	CD373397	CD373397 UI-R-GRO-
83	23.8	74.4	788	10	DT900112	DT900112 1479196 M
84	23.8	74.4	792	4	CB31634	CB31634 AGENCOURT
85	23.8	74.4	808	4	CB316171	CB316171 AGENCOURT
86	23.8	74.4	824	4	CB314729	CB314729 AGENCOURT
87	23.8	74.4	824	8	CO388606	CO388606 AGENCOURT
88	23.8	74.4	846	8	CO398822	CO398822 AGENCOURT
89	23.8	74.4	855	8	CO388172	CO388172 AGENCOURT
90	23.8	74.4	875	5	CF111188	CF111188 Shultzomi
91	23.6	73.8	216	5	CF527423	CF527423 cu-C-21
92	23.6	73.8	708	9	DM886035	DM886035 lat34a02.

969	20.8	65.0	368	4	BY106970	BY106970	BY106970
970	20.8	65.0	368	4	BY348791	BY348791	BY348791
971	20.8	65.0	368	4	BY787722	BY787722	BY787722
972	20.8	65.0	369	4	BY117245	BY117245	BY117245
973	20.8	65.0	369	4	BY773233	BY773233	BY773233
974	20.8	65.0	369	10	DT912883	DT912883	DT912883
975	20.8	65.0	370	4	BY042099	BY042099	BY042099
976	20.8	65.0	370	4	BY042810	BY042810	BY042810
977	20.8	65.0	370	4	BY150531	BY150531	BY150531
978	20.8	65.0	371	4	BY107183	BY107183	BY107183
979	20.8	65.0	371	4	BY135639	BY135639	BY135639
980	20.8	65.0	372	4	BY041240	BY041240	BY041240
981	20.8	65.0	372	4	BY042811	BY042811	BY042811
982	20.8	65.0	372	4	BY767654	BY767654	BY767654
983	20.8	65.0	373	1	AA210569	AA210569	mo86612.x
984	20.8	65.0	373	4	BY041249	BY041249	BY041249
985	20.8	65.0	373	4	BY054024	BY054024	BY054024
986	20.8	65.0	373	5	CJ087274	CJ087274	CJ087274
987	20.8	65.0	374	4	BY055334	BY055334	BY055334
988	20.8	65.0	374	4	BY069567	BY069567	BY069567
989	20.8	65.0	375	4	BY035979	BY035979	BY035979
990	20.8	65.0	375	4	BY057817	BY057817	BY057817
991	20.8	65.0	375	4	BY307200	BY307200	BY307200
992	20.8	65.0	375	4	BY767769	BY767769	BY767769
993	20.8	65.0	375	5	CJ086829	CJ086829	CJ086829
994	20.8	65.0	376	4	BY038584	BY038584	BY038584
995	20.8	65.0	376	4	BY132354	BY132354	BY132354
996	20.8	65.0	376	5	CJ077638	CJ077638	CJ077638
c 997	20.8	65.0	377	2	BM519390	12-D11-RK	BM519390
998	20.8	65.0	377	4	BY050397	BY050397	BY050397
999	20.8	65.0	377	4	BY107145	BY107145	BY107145
1000	20.8	65.0	377	4	BY352533	BY352533	BY352533

ALIGNMENTS

RESULT 1
 DM414483 316 bp mRNA linear EST 14-JAN-2006
 LOCUS HMG8015213 Human liver regeneration after partial hepatectomy Homo
 DEFINITION sapiens cDNA, mRNA sequence.

ACCESSION DM414483.1 GI:84916039

VERSION EST.

KEYWORDS Homo sapiens (human)

SOURCE

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

AUTHORS Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

JOURNAL Homidae; Homo.

COMMENT 1 (bases 1 to 316)

Xu, C.S.

liver regeneration after PH

unpublished (2003)

Contact: Cun-Shun Xu

Henan Bioengineering Key Lab

No. 148 Jianshe Road, Xinxiang City, P.R.China

Tel: 00863733328084

Fax: 00863733326524

Email: xucsx263.net.

Location/Qualifiers

1. 316

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone_lib="Liver"

/clone_id="Human liver regeneration after partial

hepatectomy"

ORIGIN

Query Match 100.0%; Score 32; DB 10; Length 316;

Best Local Similarity 100.0%; Pred. No. 0.0055;

Matches	32;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
Qy	1	TTATTATAAGTGGCTGCGTCCGAAACC	32						
Db	51	TTATTATAAGTGGCTGCGTCCGAAACC	82						

RESULT 2
 DA567416 549 bp mRNA linear EST 08-NOV-2005
 LOCUS DA567416 HBLAC1 Homo sapiens cDNA clone HBLAC1000051 5', mRNA

DEFINITION DA567416 HBLAC1 Homo sapiens cDNA clone HBLAC1000051 5', mRNA

ACCESSION DA567416

VERSION DA567416.1 GI:81180149

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

AUTHORS Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Homidae; Homo.

1 (bases 1 to 549)

Kimura, K., Makatsugu, A., Suzuki, Y., Oca, T., Nishikawa, T.,

Yamashita, R., Yamamoto, J., Sekine, M., Tsuritani, K., Makaguchi, H.,

Ishii, S., Sugiyama, T., Saito, K., Isono, Y., Irie, R., Kuehida, N.,

Yoneyama, T., Otsuka, R., Kanda, K., Yokoi, T., Kondo, H., Magatsuna, M.,

Murakawa, K., Ishida, S., Ishibashi, T., Takahashi, Fujii, A.,

Tanase, T., Nagai, K., Kikuchi, H., Nakai, K., Isegai, T., and Sugano, S.

Diversification of Transcriptional Modulation: Large-scale

Identification and Characterization of Putative Alternative

Promoters of Human Genes

Genome Res. 16 (1), 55-65 (2006)

JOURNAL PUBMED

COMMENT

Contact: Takao Isegai

FLJ Project (HRI Team)

Helix Research Institute

2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan

Tel: 81-438-52-3975

Fax: 81-438-52-3986

Email: flj-cdna@nifty.com

NEDO human cDNA project (New Energy and Industrial Technology

Developmental Organization, Japan); cDNA library construction;

Helix Research Institute (HRI), 5'-end one pass sequencing; HRI,

Research Association for Biotechnology (RAB) and Biotechnology

Center, National Institute of Technology and Evaluation; 3'-end one

pass sequencing; RAB.

Location/Qualifiers

1. 549

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="HBLAC1000051"

/cell_type="Hela cells"

/cell_line="Hela"

/clone_lib="HBLAC1"

/note="Vector: pME18SFLJ3"

ORIGIN

Query Match 100.0%; Score 32; DB 9; Length 549;

Best Local Similarity 100.0%; Pred. No. 0.006;

Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TTATTATAAGTGGCTGCGTCCGAAACC 32

Db 1 TTATTATAAGTGGCTGCGTCCGAAACC 32

RESULT 3

BE543005 890 bp mRNA linear EST 09-AUG-2000

LOCUS 601068927F1 NIH_MGC_12 Homo sapiens cDNA clone IMAGE:3455214 5',

DEFINITION mRNA sequence.

ACCESSION BE543005

VERSION BE543005.1 GI:9771650

KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae; Homo.
TITLE NIH-MGC http://mgs.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapsb@imail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/MLML at: http://image.llnl.gov
 Plate: L14M6441 row: e column: 07
 High quality sequence stop: 542.
 Location/Qualifiers
 1. 890
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:3455214"
 /tissue_type="cervical carcinoma cell line"
 /lab_host="DH10B"
 /clone_id="NH MGC 12"
 /note="Organ: cervix; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: oligo dt. Average insert size 1.4 kb. Library prepared by Life Technologies."

ORIGIN
 Query Match 100.0%; Score 32; DB 7; Length 890;
 Best Local Similarity 100.0%; Pred. No. 0.0065;
 Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY
 1 TTATTATAAGTCTCGCTGCGTGCAGAAACC 32
 268 TTATTATAAGTCTCGCTGCGTGCAGAAACC 299

Db
 268 TTATTATAAGTCTCGCTGCGTGCAGAAACC 299

RESULT 4
LOCUS BF911801 443 bp mRNA linear EST 18-JAN-2001
DEFINITION MR4-UT0051-101100-001-a10 UT0051 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF911801
VERSION BF911801.1 GI:12303259
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae; Homo.
 1 (bases 1 to 443)
 Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W.Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bata,G.S., Simpson,D.H., Brunetti,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.U. and Simpson,A.J.
 Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
 Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

TITLE
 Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil

JOURNAL
PUBMED 10737800
COMMENT

Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the PAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?l=MR4ct2-MR4-UT0051-101100-001-a10ct3=2000-11-10&t4=1)
 Seq primer: puc 18 forward
 High quality sequence start: 21
 High quality sequence stop: 443.
 Location/Qualifiers
 1. 443
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /dev_stage="Adult"
 /clone_id="UT0051"
 /note="Organ: uterus tumor; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

ORIGIN
 Query Match 80.0%; Score 25.6; DB 2; Length 443;
 Best Local Similarity 87.5%; Pred. No. 3.8;
 Matches 26; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY
 1 TTATTATAAGTCTCGCTGCGTGCAGAAACC 32
 141 TTATTATAAGTCTCGCTGCGTGCAGAAACC 110

Db
 141 TTATTATAAGTCTCGCTGCGTGCAGAAACC 110

RESULT 5
LOCUS DW445957 219 bp mRNA linear EST 14-JAN-2006
DEFINITION HHAGE043687 Human liver regeneration after partial hepatectomy Homo sapiens cDNA, mRNA sequence.
ACCESSION DW445957
VERSION DW445957.1 GI:84947513
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae; Homo.
 1 (bases 1 to 219)
 Xu,C.S.
 Liver regeneration after PH
 Unpublished (2003)
 Contact: Cun-Shuan Xu
 Henan Bioengineering Key Lab
 Henan Normal University
 No. 148 Jianshe Road, Xinxiang City, P.R.China
 Tel: 0086373328084
 Fax: 0086373326524
 Email: xucs@x263.net.

TITLE
 Location/Qualifiers
 1. 219
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /tissue_type="liver"
 /clone_id="Human liver regeneration after partial hepatectomy"

JOURNAL
PUBMED

COMMENT

ORIGIN
 Query Match 75.6%; Score 24.2; DB 10; Length 219;
 Best Local Similarity 89.7%; Pred. No. 14;
 Matches 26; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY	4	TTAATTAAGTGCTGCGGTGCCAGAAACC	32
LOCUS	DM443699		
DEFINITION	HMAGE041429 Human liver regeneration after partial hepatectomy Homo sapiens cDNA, mRNA sequence.		
ACCESSION	DM443699		
VERSION	DM443699.1	GI:84945255	
KEYWORDS	EST.		
SOURCE	Homo sapiens (human)		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo.		
REFERENCE	Xu,C.S.		
AUTHORS	Liver regeneration after PH		
TITLE	Unpublished (2003)		
JOURNAL	Contact: Cun-Shuan Xu		
COMMENT	Henan Bioengineering Key Lab Hanan Normal University No. 148 Jianshe Road, Xinxiang City, P.R.China Tel.: 0086373328084 Fax: 00863733326524 Email: xucs@x263.net..		
FEATURES	Location/Qualifiers		
SOURCE	1..229		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/tissue_type="liver"		
	/clone_lib="Human liver regeneration after partial hepatectomy"		
ORIGIN			
Query Match	75.6%; Score 24.2; DB 10; Length 229;		
Best Local Similarity	89.7%; Pred.No.14;		
Matches	26; Conservative 0; Mismatches 3; Indels 0; Gaps 0;		
OY	4	TTAATTAAGTGCTGCGGTGCCAGAAACC	32
Dn	139	TTACAGAGTGTCTGCGGTGCCAGAAACC	167
RESULT 7			
LOCUS	DM424800	287 bp	mRNA EST 14-JAN-2006
DEFINITION	HHAGE04550 Human liver regeneration after partial hepatectomy Homo sapiens cDNA, mRNA sequence.		
ACCESSION	DM424800		
VERSION	DM424800.1	GI:84926356	
KEYWORDS	EST.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo.		
REFERENCE	Xu,C.S.		
AUTHORS	Liver regeneration after PH		
TITLE	Unpublished (2003)		
JOURNAL	Contact: Cun-Shuan Xu		
COMMENT	Henan Bioengineering Key Lab Hanan Normal University No. 148 Jianshe Road, Xinxiang City, P.R.China Tel.: 00863733328084 Fax: 00863733326524 Email: xucs@x263.net.		
FEATURES	Location/Qualifiers		

source	1..289	/organism="Homo sapiens"
		/mol_type="mRNA"
		/db_xref="taxon:9606"
		/tissue_type="liver"
		/clone_idb="Human liver regeneration after partial hepatectomy"
ORIGIN		
Query Match	75.6%;	Score 24.2; DB 10; Length 289;
Best Local Similarity	89.7%;	Pred.No.15;
Matches	26; Conservative	0; Mismatches 3; Indels 0; Gaps 0;
Qy	4	TTAATTAAGTGCCTGCGGTGCCAGAAACC 32
Db	142	TTACAGAGTGCTCGCGGTGCCAGAAACC 168
RESULT 8		
LOCUS	DM424288	289 bp mRNA linear EST 14-JAN-2006
DEFINITION	HHAGE024018 Human liver regeneration after partial hepatectomy Homo	
ACCESSION	sapiens CDNA, mRNA sequence.	
VERSION	DM424288.1	GI:84925844
KEYWORDS	EST.	
SOURCE	Homo sapiens (human)	
ORGANISM	Homo sapiens	
REFERENCE	Eukaryota, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
AUTHORS	Mammalia; Eutheria; Euarchontoglires; Primates; Catherhini;	
TITLE	Hominidae; Homo.	
JOURNAL	1 (bases 1 to 289)	
COMMENT	Liver regeneration after PH	
	Unpublished (2003)	
	Contact: Cun-Shuan Xu	
	Henan Bioengineering Key Lab	
	Henan Normal University	
	No. 148 Jianshe Road, Xinxiang City, P.R.China	
	Tel: 0086373326084	
	Fax: 0086373326524	
	Email: xucsx@x263.net.	
FEATURES		
source	location/Qualifiers	
	1..289	
	/organism="Homo sapiens"	
	/mol_type="mRNA"	
	/db_xref="taxon:9606"	
	/tissue_type="liver"	
	/clone_idb="Human liver regeneration after partial hepatectomy"	
ORIGIN		
Query Match	75.6%;	Score 24.2; DB 10; Length 289;
Best Local Similarity	89.7%;	Pred.No.15;
Matches	26; Conservative	0; Mismatches 3; Indels 0; Gaps 0;
Qy	4	TTAATTAAGTGCCTGCGGTGCCAGAAACC 32
Db	142	TTACAGAGTGCTCGCGGTGCCAGAAACC 170
RESULT 9		
LOCUS	DM422532	294 bp mRNA linear EST 14-JAN-2006
DEFINITION	HHAGE022262 Human liver regeneration after partial hepatectomy Homo	
ACCESSION	sapiens cDNA, mRNA sequence.	
VERSION	DM422532	GI:84924088
KEYWORDS	EST.	
SOURCE	Homo sapiens (human)	
ORGANISM	Homo sapiens	
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
	Mammalia; Eutheria; Euarchontoglires; Primates; Catherhini;	

REFERENCE 1 (bases 1 to 294)
 AUTHORS Xu, C.S.
 TITLE Liver regeneration after PH
 JOURNAL Unpublished (2003)
 COMMENT Contact: Cun-Shuan Xu
 Henan Bioengineering Key Lab
 Henan Normal University
 No. 148 Jianshe Road, Xinxiang City, P.R.China
 Tel: 0086373328084
 Fax: 0086373326524
 Email: xucse@263.net.

FEATURES
 source
 1..294
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /tissue_type="liver"
 /clone_lib="Human liver regeneration after partial
 hepatectomy"

ORIGIN
 Query Match 75.6%; Score 24.2; DB 10; Length 294;
 Best Local Similarity 89.7%; Pred. No. 15;
 Matches 26; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 TTATAAGGTGCTCGCGTCCAGAAACC 32
 |||
 139 TTACAGAGGTGCTCGCGTCCAGAAACC 167

RESULT 10
 LOCUS DM415682 312 bp mRNA linear EST 14-JAN-2006
 DEFINITION H14G5016412 Human liver regeneration after partial hepatectomy Homo
 sapiens cDNA, mRNA sequence.
 ACCESSION DM415682
 VERSION DM415682.1 GI:84917238
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homnidae; Homo.
 1 (bases 1 to 312)
 Xu, C.S.
 Liver regeneration after PH
 Unpublished (2003)
 Contact: Cun-Shuan Xu
 Henan Bioengineering Key Lab
 Henan Normal University
 No. 148 Jianshe Road, Xinxiang City, P.R.China
 Tel: 0086373328084
 Fax: 0086373326524
 Email: xucse@263.net.

FEATURES
 source
 1..312
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /tissue_type="liver"
 /clone_lib="Human liver regeneration after partial
 hepatectomy"

ORIGIN
 Query Match 75.6%; Score 24.2; DB 10; Length 312;
 Best Local Similarity 89.7%; Pred. No. 15;
 Matches 26; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 TTATAAGGTGCTCGCGTCCAGAAACC 32
 |||
 142 TTACAGAGGTGCTCGCGTCCAGAAACC 170

RESULT 11
 LOCUS BP912441 405 bp mRNA linear EST 18-JAN-2001
 DEFINITION IL2-UT0073-161100-246-G12 UT0073 Homo sapiens cDNA, mRNA sequence.
 BP912441
 VERSION BP912441.1 GI:12303899
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homnidae; Homo.
 1 (bases 1 to 405)
 Dias Neto, R., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
 Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.P.,
 Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H.,
 Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V.,
 O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
 Simpson, A.J.

REFERENCE
 AUTHORS
 TITLE Shotgun sequencing of the human transcriptome with ORF expressed
 sequence tags
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 PUBMED 10737800
 COMMENT Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome
 Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=IL2ct=IL2-UT0073-
 161100-246-G12ct=2000-11-16ct=1)
 Seq primer: puc 18 forward
 High quality sequence stop: 404.

FEATURES
 source
 1..405
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /dev_stage="adult"
 /clone_lib="UT0073"
 /note="Organ: uterus tumor; Vector: puc18, Site_1: SmaI;
 Site_2: SmaI; A mini-library was made by cloning products
 derived from ORESTES PCR (U.S. Letters Patent application
 No. 196,716 - Ludwig Institute for Cancer Research)
 profiles into the puc 18 vector. Reverse transcription of
 tissue mRNA and cDNA amplification were performed under
 low stringency conditions."

ORIGIN
 Query Match 75.6%; Score 24.2; DB 2; Length 405;
 Best Local Similarity 89.7%; Pred. No. 15;
 Matches 26; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 TTATAAGGTGCTCGCGTCCAGAAACC 32
 |||
 323 TTACAGAGGTGCTCGCGTCCAGAAACC 351

RESULT 12
 LOCUS BP912537 405 bp mRNA linear EST 18-JAN-2001
 DEFINITION IL2-UT0073-191100-246-G12 UT0073 Homo sapiens cDNA, mRNA sequence.
 BP912537
 VERSION BP912537.1 GI:12303995
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE 1 (bases 1 to 405)
 AUTHORS Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homiidae; Homo.
 Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
 Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
 Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H.,
 Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V.,
 O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
 Simpson, A.J.
 Shotgun sequencing of the human transcriptome with ORF expressed
 sequence tags
 Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 10737800
 CONTACT: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome
 Project. This entry can be seen in the following URL:
 (http://www.ludwig.org.br/scripts/gethtml2.pl?l=IL2&t=IL2-UT0073-
 191100-246-g12&t3=2000-11-19&t4=1)
 Seq primer: puc 18 forward
 High quality sequence stop: 404.
 Location/Qualifiers
 1..405
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /dev_stage="Adult"
 /clone_lib="UT0073"
 /note="Organ: uterus tumor; Vector: puc18; Site_1: SmaI;
 Site_2: SmaI; A mini-library was made by cloning products
 derived from ORBSTS PCR (U.S. Letters Patent application
 No. 196,716 - Ludwig Institute for Cancer Research)
 profiles into the puc 18 vector. Reverse transcription of
 tissue mRNA and cDNA amplification were performed under
 low stringency conditions."
 ORIGIN
 Query Match 75.6%; Score 24.2; DB 2; Length 405;
 Best Local Similarity 89.7%; Pred. No. 15;
 Matches 26; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 4 TTATTAAGTGCTGCGGTGCCGGAACC 32
 Db 323 TTACAGAGTGCTGCGGTGCCGGAACC 351
 RESULT 13
 LOCUS DW419176 407 bp mRNA linear EST 14-JAN-2006
 DEFINITION HNA6001906 Human liver regeneration after partial hepatectomy Homo
 sapiens cDNA, mRNA sequence.
 DW419176
 ACCESSION DW419176.1 GI:84920732
 VERSION
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homiidae; Homo.
 1 (bases 1 to 407)
 Xu, C.S.
 Liver regeneration after PH
 Unpublished (2003)
 Contact: Cun-Shuan Xu
 Henan Bioengineering Key Lab
 Henan Normal University
 No. 148 Jianshe Road, Xinxiang City, P.R.China

FEATURES
 source Location/Qualifiers
 1..407
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /dev_stage="Adult"
 /clone_lib="UT0073"
 /note="Organ: uterus tumor; Vector: puc18; Site_1: SmaI;
 Site_2: SmaI; A mini-library was made by cloning products
 derived from ORBSTS PCR (U.S. Letters Patent application
 No. 196,716 - Ludwig Institute for Cancer Research)
 profiles into the puc 18 vector. Reverse transcription of
 tissue mRNA and cDNA amplification were performed under
 low stringency conditions."
 ORIGIN
 Query Match 75.6%; Score 24.2; DB 10; Length 407;
 Best Local Similarity 89.7%; Pred. No. 15;
 Matches 26; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 4 TTATTAAGTGCTGCGGTGCCGGAACC 32
 Db 140 TTACAGAGTGCTGCGGTGCCGGAACC 168
 RESULT 14
 LOCUS BF912279 410 bp mRNA linear EST 18-JAN-2001
 DEFINITION IL2-UT0073-131100-245-G08 UT0073 Homo sapiens cDNA, mRNA sequence.
 BF912279
 ACCESSION BF912279
 VERSION BF912279.1 GI:12303737
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homiidae; Homo.
 1 (bases 1 to 410)
 Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
 Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
 Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H.,
 Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V.,
 O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
 Simpson, A.J.
 Shotgun sequencing of the human transcriptome with ORF expressed
 sequence tags
 Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 10737800
 CONTACT: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome
 Project. This entry can be seen in the following URL:
 (http://www.ludwig.org.br/scripts/gethtml2.pl?l=IL2&t=IL2-UT0073-
 131100-245-g08&t3=2000-11-13&t4=1)
 Seq primer: puc 18 forward
 High quality sequence stop: 409.
 Location/Qualifiers
 1..410
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /dev_stage="Adult"
 /clone_lib="UT0073"
 /note="Organ: uterus tumor; Vector: puc18; Site_1: SmaI;
 Site_2: SmaI; A mini-library was made by cloning products
 derived from ORBSTS PCR (U.S. Letters Patent application
 No. 196,716 - Ludwig Institute for Cancer Research)
 profiles into the puc 18 vector. Reverse transcription of
 tissue mRNA and cDNA amplification were performed under
 low stringency conditions."

ORIGIN

Query Match 75.6%; Score 24.2; DB 2; Length 410;
Best Local Similarity 89.7%; Pred. No. 15;
Matches 26; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 4 TTATAAGTGCCTGCGTGCAGAAACC 32
|||||
324 TTACAGAGTGCTCGCTGCAGAAACC 352

RESULT 15
AM675319 420 bp mRNA linear EST 11-APR-2000
LOCUS
DEFINITION
bp44c01.y1 NIH-MGC 12 Homo sapiens CDNA clone IMAGE:2986080 5'
similar to SW:V66_HPV18 P06463 B6 PROTEIN.; mRNA sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AM675319
AM675319
AM675319.1 GI:7540554
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Homo.
1 (bases 1 to 420)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strauberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
image.lnl.gov/image/html/resources.shtml
Seq primer: -40RP from Glbco.

FEATURES
Source
Location/Qualifiers
1..420
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2986080"
/tissue_type="cervical carcinoma cell line"
/lab_host="DH10B"
/clone_lib="NIH-MGC 12"
/note="Organ: cervix; Vector: PCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: oligo dt.
Average insert size 1.4 kb. Library prepared by Life
Technologies."

ORIGIN

Query Match 75.6%; Score 24.2; DB 7; Length 420;
Best Local Similarity 89.7%; Pred. No. 15;
Matches 26; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 4 TTATAAGTGCCTGCGTGCAGAAACC 32
|||||
107 TTACAGAGTGCTCGCTGCAGAAACC 135

RESULT 16
BE018031 509 bp mRNA linear EST 06-JUN-2000
LOCUS
DEFINITION
bp73c04.y1 NIH-MGC 12 Homo sapiens CDNA clone IMAGE:3048006 5'
similar to SW:V66_HPV18 P06463 B6 PROTEIN.; mRNA sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

BE018031
BE018031
BE018031.1 GI:8277975
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Homo.
1 (bases 1 to 509)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strauberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
image.lnl.gov/image/html/resources.shtml
Seq primer: -40RP from Glbco
High quality sequence stop: 441.
Location/Qualifiers
1..509
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3048006"
/tissue_type="cervical carcinoma cell line"
/lab_host="DH10B"
/clone_lib="NIH-MGC 12"
/note="Organ: cervix; Vector: PCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: oligo dt.
Average insert size 1.4 kb. Library prepared by Life
Technologies."

ORIGIN

Query Match 75.6%; Score 24.2; DB 7; Length 509;
Best Local Similarity 89.7%; Pred. No. 16;
Matches 26; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 4 TTATAAGTGCCTGCGTGCAGAAACC 32
|||||
94 TTACAGAGTGCTCGCTGCAGAAACC 122

RESULT 17
B1260198 742 bp mRNA linear EST 17-JUL-2001
LOCUS
DEFINITION
60297024F1 NIH-MGC 12 Homo sapiens CDNA clone IMAGE:5109562 5',
mRNA sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

B1260198
B1260198
B1260198.1 GI:14818276
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Homo.
1 (bases 1 to 742)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strauberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: Incyte Genomics, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
Plate: LAM11265 row: h column: 11
High quality sequence stop: 738.
Location/Qualifiers
1..742
/organism="Homo sapiens"
/mol_type="mRNA"

FEATURES
Source
Location/Qualifiers

/db_xref="taxon:9606"
/clone="IMAGE:5109562"
/tissue_type="cervical carcinoma cell line"
/lab_host="DH10B"
/clone_1lb="NIH_MGC_12"
/note="Organ: cervix; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.4 kb. Library prepared by Life
Technologies."

ORIGIN

Query Match 75.6%; Score 24.2; DB 2; Length 742;

Best Local Similarity 89.7%; Pred. No. 17;
Matches 26; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 TTATATAGTGCTGCGGTGCCGAGAAACC 32
DB 302 TTACAGAGTGCTGCGGTGCCGAGAAACC 330

RESULT 18
BI333226 748 bp mRNA linear EST 30-JUL-2001
LOCUS 602996552F1 NIH_MGC_12 Homo sapiens cDNA clone IMAGE:5138749 5',
DEFINITION mRNA sequence.
ACCESSION BI333226
VERSION BI333226.1 GI:15017883
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLES National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cga@bs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: Incyte Genomics, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LINL at:
http://image.lnl.gov
Plate: LHAM11341 row: h column: 14
High quality sequence stop: 723.
Location/Qualifiers

FEATURES
source 1..748
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5138749"
/tissue_type="cervical carcinoma cell line"
/lab_host="DH10B"
/clone_1lb="NIH_MGC_12"
/note="Organ: cervix; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.4 kb. Library prepared by Life
Technologies."

ORIGIN

Query Match 75.6%; Score 24.2; DB 2; Length 748;

Best Local Similarity 89.7%; Pred. No. 17;
Matches 26; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 TTATATAGTGCTGCGGTGCCGAGAAACC 32
DB 103 TTACAGAGTGCTGCGGTGCCGAGAAACC 131

RESULT 19

BI222960
LOCUS BI222960 758 bp mRNA linear EST 11-JUL-2001
DEFINITION 602943447F1 NIH_MGC_12 Homo sapiens cDNA clone IMAGE:5106611 5',
mRNA sequence.

ACCESSION BI222960
VERSION BI222960.1 GI:14676404
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLES National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cga@bs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: Incyte Genomics, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LINL at:
http://image.lnl.gov
Plate: LHAM11257 row: m column: 12
High quality sequence stop: 717.
Location/Qualifiers

FEATURES
source 1..758
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5106611"
/tissue_type="cervical carcinoma cell line"
/lab_host="DH10B"
/clone_1lb="NIH_MGC_12"
/note="Organ: cervix; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.4 kb. Library prepared by Life
Technologies."

ORIGIN

Query Match 75.6%; Score 24.2; DB 2; Length 758;

Best Local Similarity 89.7%; Pred. No. 17;
Matches 26; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 TTATATAGTGCTGCGGTGCCGAGAAACC 32
DB 107 TTACAGAGTGCTGCGGTGCCGAGAAACC 135

RESULT 20
BI255152 827 bp mRNA linear EST 17-JUL-2001
LOCUS 602977434F1 NIH_MGC_12 Homo sapiens cDNA clone IMAGE:5122136 5',
mRNA sequence.
ACCESSION BI255152
VERSION BI255152.1 GI:14808275
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLES National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cga@bs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: Incyte Genomics, Inc.

DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/INL at:
<http://image.lnl.gov>
 Plate: LHM11298 row: d column: 09
 High quality sequence stop: 742.
 Location/Qualifiers

FEATURES

source

1.827
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5122136"
 /tissue_type="cervical carcinoma cell line"
 /lab_host="DH10B"
 /clone_lib="NIH_MGC_12"
 /note="Organ: cervix; Vector: pCMV-SPORT6; Site 1: NotI;
 Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
 Average insert size 1.4 kb. Library prepared by Life
 Technologies."

ORIGIN

Query Match 75.6%; Score 24.2; DB 2; Length 827;
 Best Local Similarity 89.7%; Pred. No. 17;
 Matches 26; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 TTATAAGTGCTGCGTGCAGAAACC 32
 DB 100 TTACAGAGTGCTGCGTGCAGAAACC 128

RESULT 21
 BI223045 833 bp mRNA linear EST 11-JUL-2001
 LOCUS 60294338.F1 NIH_MGC_12 Homo sapiens cDNA clone IMAGE:5106332 5',
 DEFINITION mRNA sequence.
 ACCESSION BI223045
 VERSION BI223045.1 GI:14676489
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homiidae; Homo.
 1 (bases 1 to 833)
 NIH-MGC <http://mgs.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgabs-remail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Library Arrayed by: Incyte Genomics, Inc.
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/INL at:
<http://image.lnl.gov>
 Plate: LHM11257 row: a column: 21
 High quality sequence stop: 803.
 Location/Qualifiers

1.833
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5106332"
 /tissue_type="cervical carcinoma cell line"
 /lab_host="DH10B"
 /clone_lib="NIH_MGC_12"
 /note="Organ: cervix; Vector: pCMV-SPORT6; Site 1: NotI;
 Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
 Average insert size 1.4 kb. Library prepared by Life
 Technologies."

FEATURES

source

1.833
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5106332"
 /tissue_type="cervical carcinoma cell line"
 /lab_host="DH10B"
 /clone_lib="NIH_MGC_12"
 /note="Organ: cervix; Vector: pCMV-SPORT6; Site 1: NotI;
 Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
 Average insert size 1.4 kb. Library prepared by Life
 Technologies."

ORIGIN

Query Match 75.6%; Score 24.2; DB 2; Length 833;
 Best Local Similarity 89.7%; Pred. No. 17;
 Matches 26; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 TTATAAGTGCTGCGTGCAGAAACC 32
 DB 33 TTACAGAGTGCTGCGTGCAGAAACC 61

RESULT 22

BX400902 833 bp mRNA linear EST 28-APR-2004
 LOCUS BX400902 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens
 DEFINITION cDNA clone CS0DK004YC04 5-PRIME, mRNA sequence.
 ACCESSION BX400902 GI:46845779
 VERSION BX400902.2
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

ORIGIN

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homiidae; Homo.
 1 (bases 1 to 833)
 L.M.B., Gruber,C., Jesse,J. and Polayes,D.
 Full-length cDNA libraries and normalization
 Unpublished (2001)
 On May 13, 2003 this sequence version replaced gi:30614402.
 Contact: Genoscope
 Genoscope - Centre National de Sequencage
 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
 end enriched, double-strand cDNA was digested with Not I and cloned
 into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
 was normalized. Library was constructed by Life Technologies, a
 division of Invitrogen. This sequence belongs to sequence cluster
 4578.r

There is a virtual cDNA representing this cluster. For more
 information about this cluster and the virtual cDNA, see
<http://www.genoscope.cns.fr/cdnas=CS0DK004B02QPl&c=4578.r>.
 Location/Qualifiers

FEATURES

source

1.833
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DK004YC04"
 /cell_type="HELA CELLS COT 25-NORMALIZED"
 /cell_line="HELA"
 /note="1st strand cDNA was primed with a NotI-oligo(dT)
 primer. Five prime end enriched, double-strand cDNA was
 digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 75.6%; Score 24.2; DB 4; Length 833;
 Best Local Similarity 89.7%; Pred. No. 17;
 Matches 26; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 TTATAAGTGCTGCGTGCAGAAACC 32
 DB 104 TTACAGAGTGCTGCGTGCAGAAACC 132

RESULT 23

BBS46557 856 bp mRNA linear EST 09-AUG-2000
 LOCUS BBS46557 60107656.F1 NIH_MGC_12 Homo sapiens cDNA clone IMAGE:3462759 5',
 DEFINITION mRNA sequence.
 ACCESSION BBS46557 GI:9775202
 VERSION BBS46557.1
 KEYWORDS EST.
 SOURCE Homo sapiens (human)

ORIGIN

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Homo.
REFERENCE 1 (bases 1 to 856)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: Incyte Genomics, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
Plate: LHAM640 row: 0 column: 15
High quality sequence stop: 706.
Location/Qualifiers
1..856
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3462758"
/issue_type="cervical carcinoma cell line"
/lab_host="DH10B"
/clone_1lb="NIH_MGC_12"
/note="Organ: cervix; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.4 kb. Library prepared by Life
Technologies."

ORIGIN
Query Match 75.6%; Score 24.2; DB 7; Length 856;
Best Local Similarity 89.7%; Pred. No. 17;
Matches 26; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 4 TTATATAGTGCTGCGTGCCAGAAACC 32
Db 76 TTACAGAGTGCTGCGTGCCAGAAACC 104

RESULT 24
BS544755 922 bp mRNA linear EST 09-AUG-2000
LOCUS 601075520F1 NIH_MGC_12 Homo sapiens cDNA clone IMAGE:3461429 5',
DEFINITION mRNA sequence.
ACCESSION BS544755
VERSION BS544755.1 GI:9773400
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Homo.
REFERENCE 1 (bases 1 to 922)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: Incyte Genomics, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
Plate: LHAM6457 row: h column: 06
High quality sequence stop: 717.
Location/Qualifiers
1..922

FEATURES
source

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3461429"
/issue_type="cervical carcinoma cell line"
/lab_host="DH10B"
/clone_1lb="NIH_MGC_12"
/note="Organ: cervix; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.4 kb. Library prepared by Life
Technologies."

ORIGIN
Query Match 75.6%; Score 24.2; DB 7; Length 922;
Best Local Similarity 89.7%; Pred. No. 18;
Matches 26; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 4 TTATATAGTGCTGCGTGCCAGAAACC 32
Db 133 TTACAGAGTGCTGCGTGCCAGAAACC 161

RESULT 25
BI256241 1043 bp mRNA linear EST 17-JUL-2001
LOCUS 602975055F1 NIH_MGC_12 Homo sapiens cDNA clone IMAGE:5114316 5',
DEFINITION mRNA sequence.
ACCESSION BI256241
VERSION BI256241.1 GI:14810460
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Homo.
REFERENCE 1 (bases 1 to 1043)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: Incyte Genomics, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
Plate: LHAM1277 row: n column: 13
High quality sequence stop: 723.
Location/Qualifiers
1..1043
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5114316"
/issue_type="cervical carcinoma cell line"
/lab_host="DH10B"
/clone_1lb="NIH_MGC_12"
/note="Organ: cervix; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.4 kb. Library prepared by Life
Technologies."

ORIGIN
Query Match 75.6%; Score 24.2; DB 2; Length 1043;
Best Local Similarity 89.7%; Pred. No. 18;
Matches 26; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 4 TTATATAGTGCTGCGTGCCAGAAACC 32
Db 96 TTACAGAGTGCTGCGTGCCAGAAACC 124

RESULT 26
LOCUS BI257598 1114 bp mRNA linear EST 17-JUL-2001
DEFINITION M02967170F1 NIH_MGC_12 Homo sapiens cDNA clone IMAGE:5106952 5',
602967170F1 NIH_MGC_12 Homo sapiens cDNA clone IMAGE:5106952 5',
nDNA sequence.

ACCESSION BI257598
VERSION BI257598
KEYWORDS EST, GI:14813133
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS NIH-MGC. <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: Incyte Genomics, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/BLNT at:
<http://image.llnl.gov>
Plate: LHAM1258 row: k column: 17
High quality sequence stop: 610.
Location/Qualifiers
1..1114
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5106952"
/cissue_type="cervical carcinoma cell line"
/lab_host="DH10B"
/clone_lib="NIH_MGC_12"
/note="Organ: cervix; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.4 kb. Library prepared by Life
Technologies."

ORIGIN
Query Match 75.6%; Score 24.2; DB 2; Length 1114;
Best Local Similarity 89.7%; Pred. No. 18;
Matches 26; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Cy 4 TTATATAGGTCGCGCTGCCAGAAC 32
Db 213 TTACAGAGGTGCGCTGCCAGAAC 241

RESULT 27
LOCUS AI747399 554 bp mRNA linear EST 22-JUN-1999
DEFINITION U105H08.Y1 Sugano mouse kidney mKia Mus musculus cDNA clone
IMAGE:2065215 5' similar to gb:X57958 60S RIBOSOMAL PROTEIN L7
(HUMAN); gb:M29015 Mouse ribosomal protein L7 (MOUSE);, mRNA
sequence.

ACCESSION AI747399
VERSION AI747399
KEYWORDS EST, GI:5125663
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE
AUTHORS Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 554)
Marras, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T.,
Underwood, K., Stepec, M., Theising, B., Allen, M., Bowers, Y.,
Person, B., Swaller, T., Gibbons, M., Page, D., Harvey, N., Schurk, R.,
Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R.,

TITLE Waterston, R. and Wilson, R.
JOURNAL The WashU-NCI Mouse EST Project 1999
COMMENT Unpublished (1999)
Contact: Marra M/WashU-NCI Mouse EST Project 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:994403
Seq primer: custom primer used
High quality sequence stop: 519.
Location/Qualifiers
1..554
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL"
/db_xref="taxon:10090"
/clone="IMAGE:2065215"
/sex="Female"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="Sugano mouse kidney mKia"
/note="Organ: kidney; Vector: pM185-F13; Site 1: DraIII
(CACTGTG18); Site 2: DraIII (CACTGTG18); 1st strand cDNA
was primed with an oligo(dt) primer
[ATGCGCCCTTTTCTTTTCTTTT]; double-stranded cDNA was
ligated to a DraIII adaptor (TGTGGCCCTCTG), digested
and cloned into distinct DraIII sites of the pM185-F13
vector (5' site CACTGTG18, 3' site CACTGTG18). XhoI should
be used to isolate the cDNA insert. Site selection was
performed to exclude fragments <1.5kb. Library
constructed by Dr. Sumio Sugano (University of Tokyo
Institute of Medical Science). Custom primers for
sequencing: 5' end primer CTTCTGCTCTTAAAGCTGCG and 3' end
primer CACATCGACGCTGACAC."

ORIGIN
Query Match 75.0%; Score 24; DB 1; Length 554;
Best Local Similarity 84.4%; Pred. No. 20;
Matches 27; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Cy 1 TTATATAGGTCGCGCTGCCAGAAC 32
Db 96 TTATTAAGAGTTCGCGCTGCCAGAAC 127

RESULT 28
LOCUS AA686665 202 bp mRNA linear EST 13-JAN-1998
DEFINITION EST110976 Rat PC-12 cells, NGF-created (9 days) Rattus norvegicus
cDNA 5' and similar to Ribosomal protein L7, mRNA sequence.

ACCESSION AA686665
VERSION AA686665
KEYWORDS EST, GI:2673263
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus

REFERENCE
AUTHORS Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Rattus.
1 (bases 1 to 202)
Lee, N.H., Weinstein, K.G., Kirkness, E.F., Earle-Hughes, J.A.,
Fulcher, R.A., Marmaras, S., Glodok, A., Gocayne, J.D., Adams, M.D.,
Kerlavage, A.R., Fraser, C.M. and Venter, J.C.
Comparative expressed-sequence-tag analysis of differential gene
expression profiles in PC-12 cells before and after nerve growth
factor treatment
Proc. Natl. Acad. Sci. U.S.A. 92, 8303-8307 (1995)
Other ESTs: TC44726
Contact: Lee, NH

The Institute for Genomic Research
9712, Medical Center Drive, Rockville, MD 20850, USA
Tel: (301)-838-3529
Fax: (301)-838-0208
Email: nhlee@tigr.org
Seq primer: M13 Reverse.

FEATURES

source

```
1. .202
/organism="Rattus norvegicus"
/mol_type="RNA"
/db_xref="ATCC (thost):2005132"
/db_xref="taxon:10116"
/clone_lib="Rat PC-12 cells, NGF-treated (9 days)"
/notes="Vector: Bluescript SK-; Site 1: EcoRI; Site 2:
XhoI; poly(A)+ RNA was purified from 9-day NGF treated
PC12 cells. cDNA was constructed using an oligo-dT primer
and directionally cloned using the Lambda ZAP II Vector
Kit by Stratagene"
```

ORIGIN

Query Match 74.4%; Score 23.8; DB 1; Length 202;
Best Local Similarity 92.6%; Pred. No. 21;
Matches 25; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 6 AATPAGGTGCTCGGTGCCGGAACC 32
53 AAAAGGTTCTCGGTGCCGGAACC 79

RESULT 29 212 bp mRNA linear EST 03-APR-1998
LOCUS AA686627
DEFINITION BT110842 Rat PC-12 cells, NGF-treated (9 days) Rattus norvegicus
cDNA clone RPBN47 5' end similar to Ribosomal protein L7, mRNA
sequence.

ACCESSION AA686627
VERSION AA686627.1 GI:2673225

KEYWORDS Rattus norvegicus (Norway rat)
SOURCE Rattus norvegicus

ORGANISM Rattus norvegicus; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Rattus.

REFERENCE

AUTHORS

Lee, N.H., Weinstein, K.G., Kirkness, E.F., Earle-Hughes, J.A.,
Fulder, R.A., Marmaraz, S., Glodok, A., Gocayne, J.D., Adams, M.D.,
Kerlavage, A.R., Fraser, C.M. and Venter, J.C.
Comparative expressed-sequence-tag analysis of differential gene
expression profiles in PC-12 cells before and after nerve growth
factor treatment
Proc. Natl. Acad. Sci. U.S.A. 92, 8303-8307 (1995)

COMMENT 7667285
Other ESTs: TC44726

Contact: Lee, NH
The Institute for Genomic Research
9712, Medical Center Drive, Rockville, MD 20850, USA
Tel: (301)-838-3529
Fax: (301)-838-0208
Email: nhlee@tigr.org
Seq primer: M13 Reverse.

FEATURES

source

```
1. .212
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="ATCC (thost):2005002"
/db_xref="taxon:10116"
/clone_lib="Rat PC-12 cells, NGF-treated (9 days)"
/notes="Vector: Bluescript SK-; Site 1: EcoRI; Site 2:
XhoI; poly(A)+ RNA was purified from 9-day NGF treated
PC12 cells. cDNA was constructed using an oligo-dT primer
and directionally cloned using the Lambda ZAP II Vector
Kit by Stratagene"
```

ORIGIN

Query Match 74.4%; Score 23.8; DB 1; Length 212;
Best Local Similarity 92.6%; Pred. No. 21;
Matches 25; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 6 AATPAGGTGCTCGGTGCCGGAACC 32
36 AAAAGGTTCTCGGTGCCGGAACC 62

RESULT 30 289 bp mRNA linear EST 06-JAN-2006
LOCUS DW320563
DEFINITION LRAGB045391 Liver regeneration after partial hepatectomy Rattus
norvegicus cDNA, mRNA sequence.

ACCESSION DW320563
VERSION DW320563.1 GI:84548946

KEYWORDS

Rattus norvegicus (Norway rat)
SOURCE Rattus norvegicus

ORGANISM Rattus norvegicus; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Rattus.

REFERENCE 1 (bases 1 to 289)
XU, C.S.
Liver regeneration after PH
Unpublished (2003)
Contact: Cun-Shuan Xu
Henan Bioengineering Key Lab
Henan Normal University
No. 148 Jianshe Road, Xinxiang City, P.R.China
Tel: 0086373328084
Fax: 0086373326524
Email: xuc@x263.net.

FEATURES

source

```
1. .289
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
/clone_lib="Liver regeneration after partial hepatectomy"
```

ORIGIN

Query Match 74.4%; Score 23.8; DB 10; Length 289;
Best Local Similarity 92.6%; Pred. No. 22;
Matches 25; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 6 AATPAGGTGCTCGGTGCCGGAACC 32
82 AAAAGGTTCTCGGTGCCGGAACC 108

RESULT 31 299 bp mRNA linear EST 13-JAN-1998
LOCUS AA685796
DEFINITION BT1108189 Rat PC-12 cells, untreated Rattus norvegicus cDNA 5' end
similar to Ribosomal protein L7, mRNA sequence.

ACCESSION AA685796
VERSION AA685796.1 GI:2672394

KEYWORDS

Rattus norvegicus (Norway rat)
SOURCE Rattus norvegicus

ORGANISM Rattus norvegicus; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Rattus.

REFERENCE

AUTHORS

Lee, N.H., Weinstein, K.G., Kirkness, E.F., Earle-Hughes, J.A.,
Fulder, R.A., Marmaraz, S., Glodok, A., Gocayne, J.D., Adams, M.D.,
Kerlavage, A.R., Fraser, C.M. and Venter, J.C.
Comparative expressed-sequence-tag analysis of differential gene
expression profiles in PC-12 cells before and after nerve growth
factor treatment

JOURNAL
PUBMED
COMMENT
7667285
Other_ESTRs: TC44726
Contact: Lee, NH
The Institute for Genomic Research
9712, Medical Center Drive, Rockville, MD 20850, USA
Tel: (301)-838-3529
Fax: (301)-838-0208
Email: nhlee@tigr.org
Seq primer: M13 Reverse.
Location/Qualifiers
1. .299
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="ATCC (inhost):2002867"
/db_xref="taxon:10116"
/clone_lib="Rat PC-12 cells, untreated"
/note="Vector: pBluescript SK-; Site 1: EcoRI; Site 2: XhoI; poly(A) + RNA was purified from untreated PC12 cells cultured for 9 days. cDNA was constructed using an oligo-dT primer and directionally cloned using the Lambda ZAP II Vector Kit by Stratagene"

ORIGIN
Query Match 74.4%; Score 23.8; DB 1; Length 299;
Best Local Similarity 92.6%; Pred. No. 22;
Matches 25; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 6 AATAAGTGCTCGGTCGCCAGAAACC 32
|||
42 AAAAAGTTCCTGCGGTGCCAGAAACC 68
|||

RESULT 32
DW300318 323 bp mRNA linear EST 06-JAN-2006
LOCUS IRAGE062146 Liver regeneration after partial hepatectomy Rattus
DEFINITION norvegicus CDNA, mRNA sequence.
ACCESSION DW300318 GI:84528697
VERSION DW300318
KEYWORDS EST.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Rattus.
1 (bases 1 to 323)
Xu, C.S.
Liver regeneration after PH
Unpublished (2003)
Contact: Cun-Shuan Xu
Henan Bioengineering Key Lab
Henan Normal University
No. 148 Jianshe Road, Xinxiang City, P.R.China
Tel: 00863733328084
Fax: 00863733326524
Email: xucs@x263.net.

FEATURES
Source
1. .323
Location/Qualifiers
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
/clone_lib="Liver regeneration after partial hepatectomy"

ORIGIN
Query Match 74.4%; Score 23.8; DB 10; Length 323;
Best Local Similarity 92.6%; Pred. No. 22;
Matches 25; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 6 AATAAGTGCTCGGTCGCCAGAAACC 32
|||
199 AAAAAGTTCCTGCGGTGCCAGAAACC 173
|||

RESULT 33
AA685923 341 bp mRNA linear EST 03-APR-1998
LOCUS EST108644 Rat PC-12 cells, NGF-treated (9 days) Rattus norvegicus
DEFINITION CDNA clone RPNAD62 5' end similar to Ribosomal protein L7, mRNA
sequence.
AA685923
AA685923.1 GI:2672521
EST.
Rattus norvegicus (Norway rat)
Source
Rattus norvegicus
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Rattus.
1 (bases 1 to 341)
Lee, N.H., Weinstein, K.G., Kirkness, E.F., Earle-Hughes, J.A.,
Pulver, R.A., Marmar, S., Glodde, A., Gocayne, J.D., Adams, M.D.,
Kerlavage, A.R., Fraser, C.M. and Venter, J.C.
Comparative expressed-sequence-tag analysis of differential gene
expression profiles in PC-12 cells before and after nerve growth
factor treatment
Proc. Natl. Acad. Sci. U.S.A. 92, 8303-8307 (1995)
7667285
Other_ESTRs: TC44726
Contact: Lee, NH
The Institute for Genomic Research
9712, Medical Center Drive, Rockville, MD 20850, USA
Tel: (301)-838-3529
Fax: (301)-838-0208
Email: nhlee@tigr.org
Seq primer: M13 Reverse.
Location/Qualifiers
1. .341
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="ATCC (inhost):2003267"
/db_xref="taxon:10116"
/clone_lib="RPNAD62"
/clone_lib="Rat PC-12 cells, NGF-treated (9 days)"
/note="Vector: pBluescript SK-; Site 1: EcoRI; Site 2: XhoI; poly(A) + RNA was purified from 9-day NGF treated PC12 cells. cDNA was constructed using an oligo-dT primer and directionally cloned using the Lambda ZAP II Vector Kit by Stratagene"

ORIGIN
Query Match 74.4%; Score 23.8; DB 1; Length 341;
Best Local Similarity 92.6%; Pred. No. 22;
Matches 25; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 6 AATAAGTGCTCGGTCGCCAGAAACC 32
|||
41 AAAAAGTTCCTGCGGTGCCAGAAACC 67
|||

RESULT 34
BE127031 344 bp mRNA linear EST 15-JUN-2000
LOCUS DBPA0780 Rat Lambda ZAP Express Library Rattus norvegicus CDNA 5',
DEFINITION mRNA sequence.
BE127031
BE127031.1 GI:8545739
EST.
Rattus norvegicus (Norway rat)
Source
Rattus norvegicus
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Rattus.
1 (bases 1 to 344)
Sleeman, M.A., Murlison, J.G., Strachan, L., Kumble, K.D., Glenn, M.P.,
McGrath, A., Grierson, A., Havukkala, I., Tan, P.L.J. and Watson, J.D.

TITLE
Expressed sequence tags of cDNA clones from rat dermal papilla cells
JOURNAL
Unpublished (2000)
COMMENT
Contact: Sleeman MA
Biology
Genesis Research and Development Corporation Limited
P.O. Box 50, 1 Fox St, Parnell, Auckland, New Zealand
Tel: 0064 9 373 5600
Fax: 0064 9 373 2189
Email: m.sleeman@genesis.co.nz
Seq primer: T3 forward
High quality sequence stop: 344.
Location/Qualifiers
1. 346
/organism="Rattus norvegicus"
/mol_type="mRNA"
/strain="Dark-Agouti"
/db_xref="taxon:10116"
/tissue_type="vibrissae"
/cell_type="dermal papilla"
/clone_lib="Rat Lambda ZAP Express Library"

ORIGIN
Query Match 74.4%; Score 23.8; DB 7; Length 344;
Best Local Similarity 92.6%; Pred. No. 22;
Matches 25; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 6 AATAAGTGCTGCGGTGCCAGAAACC 32
Db 40 AAAAGGTTCTGCGGTGCCAGAAACC 66

RESULT 35
AI229881 346 bp mRNA linear EST 20-JAN-1999
LOCUS
EST226576 Normalized rat embryo, Bento Soares Rattus sp. cDNA clone
DEFINITION
Rattus norvegicus
ACCESSION
AI229881
VERSION
AI229881.1 GI:3813768
KEYWORDS
EST.
SOURCE
Rattus sp.
ORGANISM
Rattus sp.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Rattus.
1 (bases 1 to 346)
Lee, N.H., Glodok, A., Chandra, I., Mason, T.M., Quackenbush, J.,
Kerlavage, A.R., and Adams, M.D.
Rat Genome Project: Generation of a Rat EST (RST) Catalog & Rat
Gene Index
Unpublished (1998)
JOURNAL
Contact: Lee, NH
COMMENT
The Institute for Genomic Research
9712, Medical Center Drive, Rockville, MD 20850, USA
Tel: (301)-838-3529
Fax: (301)-838-0208
Email: nhlee@tigr.org
Seq primer: M13-21.
Location/Qualifiers
1. 346
/organism="Rattus sp."
/mol_type="mRNA"
/db_xref="ATCC (inhost):2036659"
/db_xref="taxon:10118"
/clone="RATC80"
/dev_stage="embryo 8, 12, 18 dpc"
/clone_lib="Normalized rat embryo, Bento Soares"
/note="Vector: pT73Pac; Site_1: EcoRI; Site_2: NotI"

ORIGIN
Query Match 74.4%; Score 23.8; DB 1; Length 346;
Best Local Similarity 92.6%; Pred. No. 23;
Matches 25; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 6 AATAAGTGCTGCGGTGCCAGAAACC 32
Db 73 AAAAGGTTCTGCGGTGCCAGAAACC 99

RESULT 36
H35039 346 bp mRNA linear EST 02-APR-1998
LOCUS
EST109541 Rat PC-12 cells, NGF-treated (9 days) Rattus norvegicus
DEFINITION
cDNA clone RPMS64 similar to Ribosomal protein L7, mRNA sequence.
H35039
ACCESSION
H35039.1 GI:980456
VERSION
H35039.1
KEYWORDS
EST.
SOURCE
Rattus norvegicus (Norway rat)
ORGANISM
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Rattus.
1 (bases 1 to 346)
Lee, N.H., Weinstein, K.G., Kirkness, E.F., Barle-Hughes, J.A.,
Pulder, R.A., Marmar, S., Glodok, A., Gocayne, J.D., Adams, M.D.,
Kerlavage, A.R., Fraser, C.M., and Venter, J.C.
Comparative expressed-sequence-tag analysis of differential gene
expression profiles in PC-12 cells before and after nerve growth
factor treatment
Proc. Natl. Acad. Sci. U.S.A. 92, 8303-8307 (1995)
7667285
Other ESTs: TC226
Contact: Lee, NH
COMMENT
The Institute for Genomic Research
9712, Medical Center Drive, Rockville, MD 20850, USA
Tel: (301)-838-3529
Fax: (301)-838-0208
Email: nhlee@tigr.org
For clone availability please contact the TIGR Database
(tbinf@tigr.org) TC (tentative Consensus) numbers represent
assemblies of ESTs.
Location/Qualifiers
1. 346
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="ATCC (inhost):2003904"
/db_xref="taxon:10116"
/clone="RPMS64"
/note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
XhoI; poly(A)+ RNA was purified from 9-day NGF treated
PC12 cells. cDNA was constructed using an oligo-dT primer
and directionally cloned using the lambda ZAP II Vector
Kit by Stratagene"

ORIGIN
Query Match 74.4%; Score 23.8; DB 10; Length 346;
Best Local Similarity 92.6%; Pred. No. 23;
Matches 25; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 6 AATAAGTGCTGCGGTGCCAGAAACC 32
Db 69 AAAAGGTTCTGCGGTGCCAGAAACC 95

RESULT 37
DM311724 347 bp mRNA linear EST 06-JAN-2006
LOCUS
LRAGE072552 Liver regeneration after partial hepatectomy Rattus
norvegicus cDNA, mRNA sequence.
DM311724
ACCESSION
DM311724.1 GI:84540107
VERSION
DM311724.1
KEYWORDS
EST.
SOURCE
Rattus norvegicus (Norway rat)
ORGANISM
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Muridae; Murinae; Rattus.
1 (bases 1 to 347)
AUTHORS Xu, C.S.
TITLE Liver regeneration after PH
JOURNAL Unpublished (2003)
COMMENT Contact: Cun-Shuan Xu
Henan Bioengineering Key Lab
Henan Normal University
No. 148 Jianshe Road, Xinxiang City, P.R.China
Tel: 0086373328084
Fax: 0086373326524
Email: xucs@x263.net.

FEATURES

source
1..347
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
/issue_type="liver"
/clone_lib="liver regeneration after partial hepatectomy"

ORIGIN
Query Match 74.4%; Score 23.8; DB 10; Length 347;
Best Local Similarity 92.6%; Pred. No. 23;
Matches 25; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 6 AATAAGTGCCTCGGTCGCGCAAGAAC 32
DB 248 AAAAAGTCTCTCGGTCGCGCAAGAAC 222

RESULT 38

DM380853

LOCUS DM380853 350 bp mRNA linear EST 14-JUN-2006
DEFINITION LRAG8025508 Liver regeneration after partial hepatectomy Rattus
norvegicus cDNA, mRNA sequence.
ACCESSION DM380853
VERSION DM380853.1 GI:84884975
KEYWORDS EST.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Muridae; Murinae; Rattus.
1 (bases 1 to 350)

AUTHORS Xu, C.S.
TITLE Liver regeneration after PH
JOURNAL Unpublished (2003)
COMMENT Contact: Cun-Shuan Xu
Henan Bioengineering Key Lab
Henan Normal University
No. 148 Jianshe Road, Xinxiang City, P.R.China
Tel: 0086373328084
Fax: 0086373326524
Email: xucs@x263.net.

FEATURES
source
1..350
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
/issue_type="liver"
/clone_lib="liver regeneration after partial hepatectomy"

ORIGIN
Query Match 74.4%; Score 23.8; DB 10; Length 350;
Best Local Similarity 92.6%; Pred. No. 23;
Matches 25; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 6 AATAAGTGCCTCGGTCGCGCAAGAAC 32
DB 74 AAAAAGTCTCTCGGTCGCGCAAGAAC 100

RESULT 39
AJ647449 359 bp mRNA linear EST 07-JUL-2004
LOCUS AJ647449 CSEORAN19 Sus scrofa cDNA clone C0003105_L12, mRNA
DEFINITION
sequence.
ACCESSION AJ647449
VERSION AJ647449.1 GI:49324294
KEYWORDS EST.
SOURCE Sus scrofa (pig)
ORGANISM Sus scrofa

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Suina; Suidae;
Sus.

1 (bases 1 to 359)

AUTHORS Anderson, S.T., Finlayson, H.A. and Archibald, A.L.
TITLE Development of cDNA and EST resources for studying reproduction and
embryo development in pigs and cattle
JOURNAL Unpublished (2004)
COMMENT Contact: Anderson SI
Genomics and Bioinformatics
Roslin Institute
Roslin, Midlothian, EH25 9PS, UNITED KINGDOM

Single pass sequencing. Bases called and trimmed with phred
v0.02425.c. Vector identified by cross match with the -minscore 20
and -mismatch 12 options. Vector: pBluescript1(KS) R. Site1: EcoRI
R. Site2: NotI 5' Seg Primer M13P Normalised library constructed
from pooled ovaries. Clones available from UK Centre for Functional
Genomics in Farm Animals, Roslin Institute, Roslin, Midlothian, UK,
EH25 9PS, www.ark-genomics.org.

FEATURES

source

1..359
/organism="Sus scrofa"
/mol_type="mRNA"
/db_xref="taxon:9823"
/clone="C0003105_L12"
/issue_type="ovary"
/clone_lib="CSEORAN19"
/note="Vector: pBluescript1(KS+); Site 1: EcoRI; Site 2:
NotI; Single pass sequencing; Normalised library
constructed from pooled ovaries"

ORIGIN

Query Match 74.4%; Score 23.8; DB 1; Length 359;
Best Local Similarity 92.6%; Pred. No. 23;
Matches 25; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 6 AATAAGTGCCTCGGTCGCGCAAGAAC 32
DB 69 AAAAAGTCTCTCGGTCGCGCAAGAAC 95

RESULT 40

BE128703

LOCUS BE128703 367 bp mRNA linear EST 15-JUN-2000
DEFINITION DEPA2453 Rat Lambda ZAP Express Library Rattus norvegicus cDNA 5',
mRNA sequence.
ACCESSION BE128703
VERSION BE128703.1 GI:8551456
KEYWORDS EST.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Muridae; Murinae; Rattus.

1 (bases 1 to 367)

AUTHORS Sleeman, M.A., Murison, J.G., Strachan, L., Kumble, K.D., Glenn, M.P.,
McGrath, A., Grierson, A., Havukkala, I., Tan, P.L.U. and Watson, J.D.
TITLE Expressed sequence tags of cDNA clones from rat dermal papilla
cells
JOURNAL Unpublished (2000)
COMMENT Contact: Sleeman MA
Biology
Genesis Research and Development Corporation Limited

P.O. Box 50, 1 Fox St, Parnell, Auckland, New Zealand
Tel: 0064 9 373 5600
Fax: 0064 9 373 2189
Email: m.sleeman@genesis.co.nz
Seq primer: T3 forward
High quality sequence stop: 367.
Location/Qualifiers

FEATURES

source

1. 367
/organism="Rattus norvegicus"
/mol_type="mRNA"
/strain="Dark-Agouti"
/db_xref="taxon:10116"
/tissue_type="vibrissae"
/cell_type="dermal papilla"
/clone_idb="Rat Lambda ZAP Express Library"

ORIGIN

Query Match 74.4%; Score 23.8; DB 7; Length 367;
Best Local Similarity 92.6%; Pred. No. 23;
Matches 25; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 6 AATAGGTGCTGCGGCGCAAGAAC 32
DB 41 AAAAGGTTCTCGGTCGCAAGAAC 67

RESULT 41

CV971013

LOCUS CV971013 386 bp mRNA linear EST 30-NOV-2004
DEFINITION LR9801012 Liver regeneration after partial hepatectomy Rattus
norvegicus cDNA, mRNA sequence.

ACCESSION

CV971013

VERSION

CV971013.1

KEYWORDS

SOURCE

ORGANISM

Rattus norvegicus (Norway rat)
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Rattus.
1 (bases 1 to 386)
XU,C.S.
Liver regeneration after PH
Unpublished (2003)
Contact: Cun-Shuan Xu
Henan Bioengineering Key Lab
Henan Normal University
No. 148 Jianshe Road, Xinxiang City, P.R.China
Tel: 0086373328084
Fax: 0086373326524
Email: xucsx263.net.

FEATURES

source

1. 386
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
/tissue_type="liver"
/clone_idb="Liver regeneration after partial hepatectomy"

ORIGIN

Query Match 74.4%; Score 23.8; DB 8; Length 386;
Best Local Similarity 92.6%; Pred. No. 23;
Matches 25; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 6 AATAGGTGCTGCGGCGCAAGAAC 32
DB 110 AAAAGGTTCTCGGTCGCAAGAAC 136

RESULT 42

DW383713

LOCUS DW383713 386 bp mRNA linear EST 14-JAN-2006
DEFINITION LR98028724 Liver regeneration after partial hepatectomy Rattus
norvegicus cDNA, mRNA sequence.

ACCESSION DW383713
VERSION DW383713.1
KEYWORDS GI:84887835
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus

REFERENCE Xu,C.S.
TITLE Liver regeneration after PH
JOURNAL Unpublished (2003)
COMMENT Contact: Cun-Shuan Xu
Henan Bioengineering Key Lab
Henan Normal University
No. 148 Jianshe Road, Xinxiang City, P.R.China
Tel: 0086373328084
Fax: 0086373326524
Email: xucsx263.net.

1. 386
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
/tissue_type="liver"
/clone_idb="Liver regeneration after partial hepatectomy"

FEATURES

source

ORIGIN

Query Match 74.4%; Score 23.8; DB 10; Length 386;
Best Local Similarity 92.6%; Pred. No. 23;
Matches 25; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 6 AATAGGTGCTGCGGCGCAAGAAC 32
DB 110 AAAAGGTTCTCGGTCGCAAGAAC 136

RESULT 43

DW383827

LOCUS DW383827 387 bp mRNA linear EST 14-JAN-2006
DEFINITION LR98028667 Liver regeneration after partial hepatectomy Rattus
norvegicus cDNA, mRNA sequence.

ACCESSION DW383827
VERSION DW383827.1
KEYWORDS GI:84887949

SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus

REFERENCE Xu,C.S.
TITLE Liver regeneration after PH
JOURNAL Unpublished (2003)
COMMENT Contact: Cun-Shuan Xu
Henan Bioengineering Key Lab
Henan Normal University
No. 148 Jianshe Road, Xinxiang City, P.R.China
Tel: 0086373328084
Fax: 0086373326524
Email: xucsx263.net.

1. 387
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
/tissue_type="liver"
/clone_idb="Liver regeneration after partial hepatectomy"

FEATURES

source

ORIGIN

Query Match 74.4%; Score 23.8; DB 10; Length 387;
Best Local Similarity 92.6%; Pred. No. 23;
Matches 25; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 6 AATPAGGTGCTGGCGTCCAGAAACC 32
 Db 115 AAAAAGTCTCTGGCGTCCAGAAACC 141

RESULT 44
 DM385499/c 398 bp mRNA linear EST 14-JAN-2006
 LOCUS LKAGB029676 Liver regeneration after partial hepatectomy Rattus
 DEFINITION norvegicus cDNA, mRNA sequence.
 ACCESSION DM385499
 VERSION DM385499.1 GI:84889621
 KEYWORDS EST.
 SOURCE Rattus norvegicus (Norway rat)
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muroidea; Muridae; Murinae; Rattus.
 1 (bases 1 to 398)
 Xu, C.S.
 Liver regeneration after PH
 Unpublished (2003)
 Contact: Cun-Shuan Xu
 Henan Bioengineering Key Lab
 Henan Normal University
 No. 148 Jianshe Road, Xinxiang City, P.R.China
 Tel: 0086373328084
 Fax: 0086373326524
 Email: xucs@x263.net.

FEATURES
 source
 1. 398
 Location/Qualifiers
 /organism="Rattus norvegicus"
 /mol_type="mRNA"
 /db_xref="taxon:10116"
 /clone_lib="Liver regeneration after partial hepatectomy"

ORIGIN
 Query Match 74.4%; Score 23.8; DB 10; Length 398;
 Best Local Similarity 92.6%; Pred. No. 23;
 Matches 25; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 6 AATPAGGTGCTGGCGTCCAGAAACC 32
 Db 235 AAAAAGTCTCTGGCGTCCAGAAACC 209

RESULT 45
 DM386589/c 401 bp mRNA linear EST 14-JAN-2006
 LOCUS LKAGB029928 Liver regeneration after partial hepatectomy Rattus
 DEFINITION norvegicus cDNA, mRNA sequence.
 ACCESSION DM386589
 VERSION DM386589.1 GI:84890711
 KEYWORDS EST.
 SOURCE Rattus norvegicus (Norway rat)
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muroidea; Muridae; Murinae; Rattus.
 1 (bases 1 to 401)
 Xu, C.S.
 Liver regeneration after PH
 Unpublished (2003)
 Contact: Cun-Shuan Xu
 Henan Bioengineering Key Lab
 Henan Normal University
 No. 148 Jianshe Road, Xinxiang City, P.R.China
 Tel: 0086373328084
 Fax: 0086373326524
 Email: xucs@x263.net.

FEATURES
 source
 Location/Qualifiers

source
 1. 401
 /organism="Rattus norvegicus"
 /mol_type="mRNA"
 /db_xref="taxon:10116"
 /clone_lib="Liver regeneration after partial hepatectomy"

ORIGIN
 Query Match 74.4%; Score 23.8; DB 10; Length 401;
 Best Local Similarity 92.6%; Pred. No. 23;
 Matches 25; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 6 AATPAGGTGCTGGCGTCCAGAAACC 32
 Db 238 AAAAAGTCTCTGGCGTCCAGAAACC 212

RESULT 46
 AA956776 415 bp mRNA linear EST 04-JUL-1999
 LOCUS UI-R-B1-fr-f-11-0-UI-g1 UI-R-B1 Rattus norvegicus cDNA clone
 DEFINITION UI-R-B1-fr-f-11-0-UI 3' similar to gi|206735|gb|M17422|AATPBL7 Rat
 ribosomal protein L7 mRNA, 5' end, mRNA sequence.
 ACCESSION AA956776
 VERSION AA956776.1 GI:4239250
 KEYWORDS EST.
 SOURCE Rattus norvegicus (Norway rat)
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muroidea; Muridae; Murinae; Rattus.
 1 (bases 1 to 415)
 Bonaldo, M.F., Lennon, G. and Soares, M.B.
 Normalization and subtraction: two approaches to facilitate gene
 discovery
 Genome Res. 6 (9), 791-806 (1996)
 8889548
 JOURNAL PUBMED
 On May 7, 1998 this sequence version replaced gi:3120471.
 Contact: Soares, MB
 Coordinated Laboratory for Computational Genomics
 University of Iowa
 375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
 Tel: 319 335 8250
 Fax: 319 335 9565
 Email: bento-soares@uiowa.edu
 The sequence tag present in the cDNA between the NotI site and the
 oligo-dT track served to identify it as a clone from the normalized
 adult 18-Day-Embryo library. cDNA library Preparation: M. Patricia
 Bonaldo, Ph.D. Clone distribution: clones will be available through
 Research Genetics This clone is also available through the
 I.M.A.G.E. Consortium at LNL (info@image.llnl.gov). IMAGE
 ID=1779946
 Seq primer: M13 Forward
 POLYA=No.

FEATURES
 source
 Location/Qualifiers
 1. 415
 /organism="Rattus norvegicus"
 /mol_type="mRNA"
 /strain="Sprague-Dawley"
 /db_xref="taxon:10116"
 /clone="UI-R-B1-fr-f-11-0-UI"
 /dev_stage="adult"
 /lab_host="DH10B (Life Technologies)"
 /clone_lib="UI-R-B1"
 /note="Vector: pT73D-PacI, Site 1: Not I, Site 2: Eco RI,
 The UI-R-B1 library is a subtracted library derived from
 the UI-R-E0 library. The UI-R-E0 library consisted of a
 mixture of individually tagged normalized libraries
 constructed from 8, 12 and 18-day embryo. The tag is a
 string of 3-5 nucleotides present between the Not I site
 and the oligo-dT track which allows identification of the
 library of origin of a clone within the mixture. The
 subtracted library (UI-R-B1) was constructed as follows:

PCR amplified cDNA inserts from a pool of UI-R-E0 clones from which 3' ESTs have been derived was used as a driver in a hybridization with the UI-R-E0 library in the form of single-stranded circles. The remaining single-stranded circles (subtracted library) was purified by hydroxyapatite column chromatography, converted to double-stranded circles and electroporated into DH10B bacteria (Life Technologies) to generate the UI-R-E1 library. This procedure has been previously described (Bonaldo, Lennon and Soares, Genome Research 6: 791-806, 1996) ¹⁰

ORIGIN

Query Match	74.4%	Score 23.8;	DB 1;	Length 415;
Best Local Similarity	92.6%	Pred. No. 23;		
Matches 25; Conservative	0;	Mismatches 2;	Indels 0;	Gaps 0;

```

Oy      6 AATAAGTGCTGCGGTGCCAGAAACC 32
          |||||
Db      90 AAAAAGTCTCGCGTGCCAGAAACC 11

```

RESULT	47
CE273504	
LOCUS	CE273504
DEFINITION	IGFBP Rat Lambda Zap 4 h light-exposed retina cDNA library Rattus norvegicus cDNA 5', mRNA sequence.
	478 bp mRNA linear EST 31-MAY-2004

VERSION	CF273504.1	GI:47825552
KEYWORDS	EST.	
SOURCE	Rattus norvegicus (Norway rat)	
ORGANISM	Rattus norvegicus	

REFERENCE
1 (bases 1 to 478)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidae; Muridae; Murinae; Rattus.

REFERENCE	AUTHORS	TITLE	JOURNAL
1 (pages 1 to 478)	Grawal, R., Stepczynski, J., Erickson, T., Kellin, R.M., Darrow, R., Basalou, L., Patterson, M.L., Organski, D.T. and Wong, P.	Specific changes in ribosomal protein gene expression during light-induced retinal degeneration	Unpublished (2003)

University of Alberta
G502 Biological Sciences Centre, Edmonton, Alberta T6G 2B9 Canada
Tel: 780 492 3590
Fax: 780 492 9234
Email: pwwong@ualberta.ca

PCR Primers
FORWARD: T3
BACKWARD: T7
Seq primer: SK.

FEATURES	Location/Qualifiers
source	1. .478

```

/organism="Rattus norvegicus"
/mol_type="mRNA"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/sex="male"
/tissue_type="retina"
/dev_stage="adult"
/clone_lib="Rat Lambda Zap 4 h light-exposed retina cDNA
library"
/note="Vector: Uni-ZAP XR; The Rat Lambda Zap 4 h
light-exposed retina cDNA library was constructed from
retinal RNA isolated from weanling male albino
Sprague-Dawley rats, which were maintained in darkness for
40 days. At 61 days of age rats were exposed to intense
visible light (490-580 nm; 1200 to 1400 lux) for 4 hours."

```

ORIGIN

Query Match	74.4%	Score 23.8;	DB 5;	Length 478;
Best Local Similarity	92.6%	Pred. No. 24;		
Matches 25;	Conservative 0;	Mismatches 2;	Indels 0;	Gaps 0;

QY	6	AATAAGTGCTGCGGTGCCAGAAACC	32
Db	111	AAAAGTTCCTGCGGTGCCAGAAACC	137

RESULT 48			
AA944687			
LOCUS	490 bp	mrna	linear
DEFINITION	EST2200186 Normalized rat embryo, Bento Soares Rattus sp. cDNA clone		
REMARKS	3' end, mRNA sequence.		

VERSION AA944687.1 GI:3104603

KEYWORDS	EST.
SOURCE	Rattus sp.
ORGANISM	Rattus sp.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidea; Muridae; Murinae; Rattus.

REFERENCE AUTHORS	1 (bases 1 to 490)
Lee, N.H., Glodok, A., Chandra, I., Mason, T.M., Quackenbush, J.	

TITLE	Kerlavage, A.R. and Adams, M.D. Rat Genome Project: Generation of a Rat EST (RESTR) Catalog & Rat
-------	---

Gene Index
Unpublished (1998)

COMMENT
Contact: Lee, NH
The Institute for Genomic Research
8713 Medical Center Drive Rockville MD 20850 USA

9/12, Medical Centre
Tel: (301) - 838-3525
Fax: (301) - 838-0208

Fax: (301)-856-0206
Email: nhleotlgr.org
Seq primer: M13-21.

FEATURES	Location/Qualifiers
source	1: 490

```

/organism="Rattus sp."
/mol_type="mRNA"
/db_xref="ATCC {inhost:2014151}"
/db_xref="taxon:10118"
/clone="RMA479"
/dev_stage="embryo 8, 12, 18 dpc"
/clone_id="Normalized rat embryo, Bento Soares"
/notes=Vector: pT73paf; Site 1: EcoRI; Site 2: NotI

```

ORIGIN

Query Match	74.4%	Score 23.8	DB 1	Length 490
Best Local Similarity	92.6%	Pred. Nc 24		
Matches 25	Conservative 0	Mismatches 2	Indels 0	Gaps 0

```
QY      6 AATAAGTGCTGCGGTGCCAGAACC 32
          ||| ||||| ||||| ||||| |||||
Db      73 AAAAAGTTCCTGCGGTGCCAGAACC 99
```

RESULT 49
AJ647149

LOCUS	AJ647149	498 bp	mRNA	linear	EST 07-JUL-2004
DEFINITION	AJ647149	CSEQRAN19 Sus scrofa cDNA clone C0003104_N19, mRNA			

sequence.	
ACCESSION	AJ647149

VERSION AJ647149.1 GI:49323994
KEYWORDS EST.

SOURCE ORGANISM	Substrate: Carbohydrate
Sus scrofa (pig)	
Sus scrofa	

Eukaryota; Metazoa; Chordata; Crustacea; Mammalia; Eutheria; Laurasiatheria

REFERENCE 1 (bases 1 to 498)
AUTHORS Anderson S T Pin]avson H A and

AUTHORS	TITLE
Mueller-Solberg, S. L., and Finkbeiner, S. R.	Development of cDNA and EST resources for the mouse embryo development in mice and rats

JOURNAL Unpublished (2004)
COMMENT Contact: Anderson SI

Genomics and Bioinformatics

Roslin Institute
Roslin, Midlothian, EH25 9PS, UNITED KINGDOM
Single pass sequencing. Bases called and trimmed with phred
v0.020425.c. Vector identified by cross match with the -minscore 20
and -mismatch 12 options. Vector:pbuescript11(KS) R. Site1: EcoRI
R. Site2: NotI 5' Seg Primer M13f Normalised library constructed
from pooled ovaries. Clones available from UK Centre for Functional
Genomics in Farm Animals, Roslin Institute, Roslin, Midlothian, UK,
EH25 9PS, www.ark-genomics.org.

FEATURES
source

1..498
Location/Qualifiers
/organism="Sus scrofa"
/mol_type="mRNA"
/db_xref="taxon:9823"
/clone="C0003104_M19"
/tissue_type="ovary"
/clone_lib="CSBORAN19"
/note="Vector: pbuescript11(KS+); Site_1: EcoRI; Site_2:
NotI; Single pass sequencing; Normalised library
constructed from pooled ovaries"

ORIGIN

Query Match 74.4%; Score 23.8; DB 1; Length 498;
Best Local Similarity 92.6%; Pred. No. 24;
Matches 25; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 6 AATPAGGTGCTGCGGTGCCAGAAACC 32
|||
DB 32 AAAAAGGTTCTGCGGTGCCAGAAACC 58

RESULT 50
AJ650336 499 bp mRNA linear EST 07-JUL-2004
LOCUS AJ650336 CSBORAN19 Sus scrofa cDNA clone C0003275_B08, mRNA
DEFINITION
AJ650336
sequence.
AJ650336
KEYWORDS AJ650336.1 GI:49327181
EST.
SOURCE Sus scrofa (pig)
ORGANISM Sus scrofa
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Suidae;
Sus

REFERENCE 1 (bases 1 to 499)
AUTHORS Anderson,S.I., Finlayson,H.A. and Archibald,A.L.
TITLE Development of cDNA and EST resources for studying reproduction and
embryo development in pigs and cattle
JOURNAL Unpublished (2004)
COMMENT Contact: Anderson SI
Genomics and Bioinformatics
Roslin Institute

Roslin, Midlothian, EH25 9PS, UNITED KINGDOM
Single pass sequencing. Bases called and trimmed with phred
v0.020425.c. Vector identified by cross match with the -minscore 20
and -mismatch 12 options. Vector:pbuescript11(KS) R. Site1: EcoRI
R. Site2: NotI 5' Seg Primer M13f Normalised library constructed
from pooled ovaries. Clones available from UK Centre for Functional
Genomics in Farm Animals, Roslin Institute, Roslin, Midlothian, UK,
EH25 9PS, www.ark-genomics.org.

FEATURES
source

1..499
Location/Qualifiers
/organism="Sus scrofa"
/mol_type="mRNA"
/db_xref="taxon:9823"
/clone="C0003275_B08"
/tissue_type="ovary"
/clone_lib="CSBORAN19"
/note="Vector: pbuescript11(KS+); Site_1: EcoRI; Site_2:
NotI; Single pass sequencing; Normalised library
constructed from pooled ovaries"

ORIGIN

Query Match 74.4%; Score 23.8; DB 1; Length 499;
Best Local Similarity 92.6%; Pred. No. 24;
Matches 25; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 6 AATPAGGTGCTGCGGTGCCAGAAACC 32
|||
DB 60 AAAAAGGTTCTGCGGTGCCAGAAACC 86

Search completed: May 24, 2006, 07:53:19
Job time : 2255.13 secs

THIS PAGE IS BLANK

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 24, 2006, 05:34:58 ; Search time 225.028 Seconds
(without alignments)
650.664 Million cell updates/sec

Title: US-10-601-913-85

Perfect score: 21

Sequence: 1 CAGGACACAGTGGCTTTGAC 21

Scoring table: IDENTITY_NUC

Searched: 5244920 seqs, 3486124231 residues

Total number of hits satisfying chosen parameters: 10489840

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Database :

N_Geneseq.8:*
1: geneseqn1980s:*
3: geneseqn1980s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*
14: geneseqn2005s:*
15: geneseqn2006s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Length	ID	Description
C 1	100.0	21	2	AAT72382	Aat72382 Human pap
C 2	100.0	21	2	AAT72381	Aat72381 Human pap
C 3	100.0	21	2	AAT72384	Aat72384 Human pap
C 4	100.0	21	2	AAT72383	Aat72383 Human pap
C 5	100.0	21	1	AAN93015	Aan93015 Probe for
C 6	100.0	21	1	AAN93015	Aan93015 Probe for
C 7	100.0	21	1	AAV39991	AAV39991 B6 oncoge
C 8	100.0	21	1	AAV39991	AAV39991 B6 oncoge
C 9	100.0	21	1	ABQ73533	ABQ73533 HPV-PTM r
C 10	100.0	21	1	ABQ73531	ABQ73531 HPV-PTM r
C 11	100.0	21	1	ADK26909	Adk26909 Human pap
C 12	100.0	21	1	ABQ76223	Abq76223 Human pap
C 13	100.0	21	1	ADG37149	Adg37149 459bp HPV
C 14	100.0	21	1	ADG37149	Adg37149 459bp HPV
C 15	100.0	21	1	ADL37114	Adl37114 Human ova
C 16	100.0	21	1	ADL37114	Adl37114 Human ova
C 17	100.0	21	1	ADL37114	Adl37114 Human ova
C 18	100.0	21	1	ADL37114	Adl37114 Human ova

C 19	21	100.0	477	14	AEA51113	Aea51113 Human pap
C 20	21	100.0	519	2	AAT31834	Aat31834 Human pap
C 21	21	100.0	543	10	ADF31984	Adf31984 Human pap
C 22	21	100.0	570	2	AAQ75470	Aaq75470 HPV16 B6/
C 23	21	100.0	598	5	ADL43498	Adl43498 Human ova
C 24	21	100.0	712	1	AAN91600	Aan91600 Partial n
C 25	21	100.0	743	5	ADL63123	Adl63123 Human ova
C 26	21	100.0	745	4	AAH72718	Aah72718 Human cer
C 27	21	100.0	747	12	ADQ44067	Adq44067 Nucleotid
C 28	21	100.0	747	12	ADQ44065	Adq44065 Nucleotid
C 29	21	100.0	747	12	ADQ44061	Adq44061 Nucleotid
C 30	21	100.0	747	12	ADQ44063	Adq44063 Nucleotid
C 31	21	100.0	747	12	ADQ44069	Adq44069 Nucleotid
C 32	21	100.0	747	12	ADQ44071	Adq44071 Nucleotid
C 33	21	100.0	768	13	ADP47004	Adp47004 Human pap
C 34	21	100.0	768	13	AEF40156	Aef40156 Human pap
C 35	21	100.0	776	2	AAT14663	Aat14663 B6/B7 reg
C 36	21	100.0	779	2	AAK89756	Aak89756 Probe seq
C 37	21	100.0	790	2	AAQ29389	Aaq29389 DNA encod
C 38	21	100.0	801	2	AAT31833	Aat31833 Human pap
C 39	21	100.0	822	2	AAK78792	Aak78792 HPV fusio
C 40	21	100.0	822	2	AAK29781	Aak29781 Prot.D1/3
C 41	21	100.0	822	14	AED52633	Aed52633 Fusion pr
C 42	21	100.0	879	2	AAK78795	Aak78795 HPV fusio
C 43	21	100.0	879	2	AAK29784	Aak29784 CLYTA-B6-
C 44	21	100.0	879	14	AED52641	Aed52641 Fusion pr
C 45	21	100.0	921	3	AAQ09477	Aaq09477 Human pap
C 46	21	100.0	939	6	AAQ5101	Aaq5101 Human pap
C 47	21	100.0	1000	2	AAQ08627	Aaq08627 HPV-16 fr
C 48	21	100.0	1005	1	AAN91784	Aan91784 DNA probe
C 49	21	100.0	1116	2	AAK78793	Aak78793 HPV fusio
C 50	21	100.0	1116	2	AAK29782	Aak29782 Prot.D1/3
C 51	21	100.0	1116	14	AED52637	Aed52637 Fusion pr
C 52	21	100.0	1173	2	AAK78797	Aak78797 HPV fusio
C 53	21	100.0	1173	2	AAK29786	Aak29786 CLYTA-B6-
C 54	21	100.0	1173	5	AED52645	Aed52645 Fusion pr
C 55	21	100.0	7840	5	AAFS5127	Aafs5127 Nucleotid
C 56	21	100.0	7902	2	AAT09847	Aat09847 Human pap
C 57	21	100.0	7904	2	AAT94724	Aat94724 Human pap
C 58	21	100.0	7904	2	AAT94723	Aat94723 Human pap
C 59	21	100.0	7904	2	AAK33881	Aak33881 HPV-16 ge
C 60	21	100.0	7904	2	AAV99946	Aav99946 Nucleotid
C 61	21	100.0	7904	2	AAV99952	Aav99952 Nucleotid
C 62	21	100.0	7904	8	ADA03127	Ada03127 HPV B6 se
C 63	21	100.0	7904	11	ADQ58335	Adq58335 Human pap
C 64	21	100.0	7904	14	ADZ45647	Adz45647 Human pap
C 65	21	95.2	20	2	AAQ10240	Aaq10240 D99 prime
C 66	20	95.2	20	3	AAA56745	Aaa56745 Human pap
C 67	20	95.2	20	6	ABK10263	Abk10263 Human pap
C 68	20	95.2	20	8	ABZ77461	Abz77461 PCR prime
C 69	20	95.2	20	10	ADF56472	Adf56472 Human pap
C 70	20	95.2	20	10	ACA62207	Ac62207 Human pap
C 71	20	95.2	20	12	ADQ42880	Adq42880 Primer of
C 72	20	95.2	20	15	AER63214	Aer63214 Human pap
C 73	20	95.2	21	10	ADD36832	Add36832 Human pap
C 74	20	95.2	21	10	ADD36832	Add36832 Human pap
C 75	20	90.5	19	6	AAI48278	Aai48278 Ribozyme
C 76	19	90.5	20	2	AAQ48549	Aaq48549 HPV B6/7
C 77	19	90.5	20	6	ABSS5744	Abss5744 Human pap
C 78	19	90.5	20	9	ACF35816	Acf35816 HPV16 B6
C 79	19	90.5	20	14	ADK26911	Adk26911 RT-PCR pr
C 80	19	90.5	29	2	AAO76065	Aao76065 hpv prime
C 81	18.4	87.6	20	2	AAQ10241	Aaq10241 Anti-D99
C 82	18.4	87.6	115	6	ABQ73534	Abq73534 HPV-PTM r
C 83	18.4	87.6	2321	5	AAAS84294	Aaas84294 DNA encod
C 84	18.4	87.6	2545	12	ADP26917	Adp26917 Human cad
C 85	18.4	87.6	2550	2	AAT85402	Aat85402 Human cad
C 86	18.4	87.6	2550	2	AAT61923	Aat61923 Full leng
C 87	18.4	87.6	2742	14	AEC83055	Aec83055 Breast ca
C 88	18.4	87.6	2929	10	ADG42064	Adg42064 Human CDH
C 89	18.4	87.6	3136	2	AAT85406	Aat85406 Rat full
C 90	18.4	87.6	3136	2	AAT61917	Aat61917 Full leng
C 91	18.4	87.6	192427	10	ADL13825	Adl13825 Osteoartrh

C 968	14.8	70.5	1017	6	ABK37597	Abk37597 DNA encod
C 969	14.8	70.5	1017	14	ADY72224	Ady72224 Human BMY
C 970	14.8	70.5	1017	14	ADY72223	Ady72223 Human BMY
C 971	14.8	70.5	1021	10	ADC86548	Adc86548 Human GPC
C 972	14.8	70.5	1039	10	ACN39428	Acn39428 Tumour-as
C 973	14.8	70.5	1120	10	ADG08782	Adg08782 Novel DNA
C 974	14.8	70.5	1131	5	AA885492	AA885492 Bacterial
C 975	14.8	70.5	1149	5	AA885492	AA885492 DNA encod
C 976	14.8	70.5	1188	2	AA252916	AA252916 Human pro
C 977	14.8	70.5	1210	5	AD162148	Ad162148 Human ova
C 978	14.8	70.5	1246	6	AB189741	Ab189741 Human pol
C 979	14.8	70.5	1247	6	ABK71649	Abk71649 Human dlc
C 980	14.8	70.5	1261	6	AB851788	Ab851788 Human mdd
C 981	14.8	70.5	1266	10	ADC90749	Adc90749 E. faeciu
C 982	14.8	70.5	1271	3	AA453355	AA453355 Arabidops
C 983	14.8	70.5	1330	10	ADC85574	Adc85574 Human GPC
C 984	14.8	70.5	1338	3	AAC34830	Aac34830 Arabidops
C 985	14.8	70.5	1360	12	AD116270	Ad116270 Human nuc
C 986	14.8	70.5	1374	13	ADX30111	Adx30111 Plant ful
C 987	14.8	70.5	1379	13	ADX30109	Adx30109 Plant ful
C 988	14.8	70.5	1458	12	ADP10589	Adp10589 Reference
C 989	14.8	70.5	1461	10	AD896111	Ad896111 Human uri
C 990	14.8	70.5	1461	10	ADK15020	Adk15020 Urinary s
C 991	14.8	70.5	1500	13	ADU05459	Adu05459 DNA encod
C 992	14.8	70.5	1512	12	ADM87151	Adm87151 Human pro
C 993	14.8	70.5	1520	6	ABK36130	Abk36130 CDNA sequ
C 994	14.8	70.5	1545	13	ADX45523	Adx45523 Plant ful
C 995	14.8	70.5	1559	4	AA199074	AA199074 Human exc
C 996	14.8	70.5	1559	5	AA163424	AA163424 Human kid
C 997	14.8	70.5	1611	6	ABQ54790	Abq54790 Human ova
C 998	14.8	70.5	1626	8	ADA06097	Ada06097 Human fat
C 999	14.8	70.5	1632	4	AAH15344	Aah15344 Human CDN
C 1000	14.8	70.5	1664	13	ADR26077	Adr26077 Breast ca

ALIGNMENTS

RESULT 1
AAT72382/c
ID AAT72382 standard; DNA; 21 BP.
AC AAT72382;
XX
DT 09-FEB-1998 (first entry)
XX
DE Human Papillomavirus Type 16 target region.
XX
KW Human Papillomavirus; probe; target region; genital cancer; HPV;
KM cervical smear; ss.
XX
OS Human papillomavirus.
XX
PN BP774518-A2.
XX
PD 21-MAY-1997.
XX
PF 15-NOV-1996; 96BP-00308264.
XX
PR 15-NOV-1995; 95US-0006854P.
XX
PA (GENP-) GEN-PROBE INC.
XX
PI Gordon P, Brentano ST, Carter NM, Hammond PW;
XX
DR WPI; 1997-274349/25.
XX
PT Probes for detection of Human Papillomavirus Type 16 and Type 18 - can
XX distinguish between Type 16 and 18, associated with genital cancers.
XX
PS Claim 8; Page 52; 70pp; English.
XX
CC Novel hybridisation assay probes have been developed comprising an

CC oligonucleotide which will hybridise under selected conditions to Human
CC Papillomavirus (HPV) Type 16 and/or 18 (but not Types 6, 11, 31, 33, 35,
CC 39, 45, 51, 52, or 58) target nucleic acids to form detectable target;
CC probe duplex. The present sequence represents a specifically claimed
CC target region. Oligonucleotides are useful to detect HPV Type 16 and/or
CC 18 in samples e.g. cervical smears, body fluid, and distinguish these
CC from other HPV variants. Papillomaviruses are small DNA viruses and HPV
CC Types 16 and 18 are associated with genital cancers. HPV PCR primers can
CC amplify HPV Type 16 and/or 18 nucleic acid in a sample. HPV Type 16
CC and/or 18 can be detected by adding a probe and detecting probe:target
CC duplex formation; target nucleic acid is optionally amplified. Type 16 or
CC 18 can be specifically detected by amplifying nucleic acids with at least
CC one specifically claimed PCR primer. For Type 18 detection, a helper
CC probe may be used. Oligonucleotides can distinguish target sequences from
CC phylogenetically close HPV Types 16 and 18 whilst available detection
CC sera react with antigens of all papillomaviruses

SQ Sequence 21 BP; 5 A; 6 C; 5 G; 5 T; 0 U; 0 Other;
Query Match 100.0%; Score 21; DB 2; Length 21;
Best Local Similarity 100.0%; Pred. No. 1;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CAGGACACAGTGGCTTTGAC 21
|||||
DB 21 CAGGACACAGTGGCTTTGAC 1

RESULT 2
AAT72381
ID AAT72381 standard; DNA; 21 BP.
AC AAT72381;
XX
DT 09-FEB-1998 (first entry)
XX
DE Human Papillomavirus Type 16 target region.
XX
KW Human Papillomavirus; probe; target region; genital cancer; HPV;
KM cervical smear; ss.
XX
OS Human papillomavirus.
XX
PN BP774518-A2.
XX
PD 21-MAY-1997.
XX
PF 15-NOV-1996; 96BP-00308264.
XX
PR 15-NOV-1995; 95US-0006854P.
XX
PA (GENP-) GEN-PROBE INC.
XX
PI Gordon P, Brentano ST, Carter NM, Hammond PW;
XX
DR WPI; 1997-274349/25.
XX
PT Probes for detection of Human Papillomavirus Type 16 and Type 18 - can
XX distinguish between Type 16 and 18, associated with genital cancers.
XX
PS Claim 8; Page 52; 70pp; English.
XX
CC Novel hybridisation assay probes have been developed comprising an
CC oligonucleotide which will hybridise under selected conditions to Human
CC Papillomavirus (HPV) Type 16 and/or 18 (but not Types 6, 11, 31, 33, 35,
CC 39, 45, 51, 52, or 58) target nucleic acids to form detectable target;
CC probe duplex. The present sequence represents a specifically claimed
CC target region. Oligonucleotides are useful to detect HPV Type 16 and/or
CC 18 in samples e.g. cervical smears, body fluid, and distinguish these
CC from other HPV variants. Papillomaviruses are small DNA viruses and HPV
CC Types 16 and 18 are associated with genital cancers. HPV PCR primers can
CC amplify HPV Type 16 and/or 18 nucleic acid in a sample. HPV Type 16
CC and/or 18 can be detected by adding a probe and detecting probe:target

CC duplex formation; target nucleic acid is optionally amplified. Type 16 or
 CC 18 can be specifically detected by amplifying nucleic acids with at least
 CC one specifically claimed PCR primer. For Type 18 detection, a helper
 CC probe may be used. Oligonucleotides can distinguish target sequences from
 CC phylogenetically close HPV Types 16 and 18 whilst available detection
 CC sera react with antigens of all papillomaviruses

XX Sequence 21 BP; 5 A; 5 C; 6 G; 5 T; 0 U; 0 Other;

Query Match 100.0%; Score 21; DB 2; Length 21;
 Best Local Similarity 100.0%; Pred. No. 1;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAGGACACAGTGGCTTTGAC 21
 1 CAGGACACAGTGGCTTTGAC 21
 Db

RESULT 3

AAT72384/C
 ID AAT72384 standard; RNA; 21 BP.

XX AAT72384;

AC 09-FEB-1998 (first entry)

XX Human Papillomavirus Type 16 target region.

DE Human Papillomavirus; probe; target region; genital cancer; HPV;

XX cervical smear; ss.

XX Human papillomavirus.

OS EP774518-A2.

XX 21-MAY-1997.

XX 15-NOV-1996; 96EP-00308264.

XX 15-NOV-1995; 95US-0006854P.

XX (GENP-) GEN-PROBE INC.

PA Gordon P, Brentano ST, Carter NM, Hammond PW;

PI WPI; 1997-274349/25.

XX Probes for detection of Human Papillomavirus Type 16 and Type 18 - can
 XX distinguish between Type 16 and 18, associated with genital cancers.

PS Claim 8; Page 53; 70pp; English.

XX Novel hybridisation assay probes have been developed comprising an
 CC oligonucleotide which will hybridise under selected conditions to Human
 CC Papillomavirus (HPV) Type 16 and/or 18 (but not Types 6, 11, 31, 33, 35,
 CC 39, 45, 51, 52, or 58) target nucleic acids to form detectable target:
 CC probe duplex. The present sequence represents a specifically claimed
 CC target region. Oligonucleotides are useful to detect HPV Type 16 and/or
 CC 18 in samples e.g. cervical smears, body fluid, and distinguish these
 CC from other HPV variants. Papillomaviruses are small DNA viruses and HPV
 CC Types 16 and 18 are associated with genital cancers. HPV PCR primers can
 CC amplify HPV Type 16 and/or 18 nucleic acid in a sample. HPV Type 16
 CC and/or 18 can be detected by adding a probe and detecting probe:target
 CC duplex formation; target nucleic acid is optionally amplified. Type 16 or
 CC 18 can be specifically detected by amplifying nucleic acids with at least
 CC one specifically claimed PCR primer. For Type 18 detection, a helper
 CC probe may be used. Oligonucleotides can distinguish target sequences from
 CC phylogenetically close HPV Types 16 and 18 whilst available detection
 CC sera react with antigens of all papillomaviruses

XX Sequence 21 BP; 5 A; 5 C; 6 G; 5 T; 0 U; 0 Other;

Query Match 100.0%; Score 21; DB 2; Length 21;

Best Local Similarity 100.0%; Pred. No. 1;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAGGACACAGTGGCTTTGAC 21
 1 CAGGACACAGTGGCTTTGAC 21
 Db

RESULT 4

AAT72383
 ID AAT72383 standard; RNA; 21 BP.

XX AAT72383;

AC 09-FEB-1998 (first entry)

XX Human Papillomavirus Type 16 target region.

DE Human Papillomavirus; probe; target region; genital cancer; HPV;

XX cervical smear; ss.

XX Human papillomavirus.

OS EP774518-A2.

XX 21-MAY-1997.

XX 15-NOV-1996; 96EP-00308264.

XX 15-NOV-1995; 95US-0006854P.

XX (GENP-) GEN-PROBE INC.

PA Gordon P, Brentano ST, Carter NM, Hammond PW;

PI WPI; 1997-274349/25.

XX Probes for detection of Human Papillomavirus Type 16 and Type 18 - can
 XX distinguish between Type 16 and 18, associated with genital cancers.

PS Claim 8; Page 53; 70pp; English.

XX Novel hybridisation assay probes have been developed comprising an
 CC oligonucleotide which will hybridise under selected conditions to Human
 CC Papillomavirus (HPV) Type 16 and/or 18 (but not Types 6, 11, 31, 33, 35,
 CC 39, 45, 51, 52, or 58) target nucleic acids to form detectable target:
 CC probe duplex. The present sequence represents a specifically claimed
 CC target region. Oligonucleotides are useful to detect HPV Type 16 and/or
 CC 18 in samples e.g. cervical smears, body fluid, and distinguish these
 CC from other HPV variants. Papillomaviruses are small DNA viruses and HPV
 CC Types 16 and 18 are associated with genital cancers. HPV PCR primers can
 CC amplify HPV Type 16 and/or 18 nucleic acid in a sample. HPV Type 16
 CC and/or 18 can be detected by adding a probe and detecting probe:target
 CC duplex formation; target nucleic acid is optionally amplified. Type 16 or
 CC 18 can be specifically detected by amplifying nucleic acids with at least
 CC one specifically claimed PCR primer. For Type 18 detection, a helper
 CC probe may be used. Oligonucleotides can distinguish target sequences from
 CC phylogenetically close HPV Types 16 and 18 whilst available detection
 CC sera react with antigens of all papillomaviruses

XX Sequence 21 BP; 5 A; 5 C; 6 G; 5 T; 0 U; 0 Other;

Query Match 100.0%; Score 21; DB 2; Length 21;
 Best Local Similarity 76.2%; Pred. No. 1;
 Matches 16; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAGGACACAGTGGCTTTGAC 21
 1 CAGGACACAGTGGCTTTGAC 21
 Db

RESULT 5

AAN93015

ID AAN93015 standard; DNA; 24 BP.
 AC AAN93015;
 XX
 XX
 DT 31-OCT-2002 (revised)
 DT 13-FEB-1990 (first entry)
 XX
 XX Probe for detecting human papillomavirus.
 KM Human papillomavirus; polynucleotide probe.
 XX
 OS Human papillomavirus.
 XX
 XX WO8905357-A.
 PN
 XX
 PD 15-JUN-1989.
 XX
 XX 06-DEC-1988; 88WO-US004355.
 PF
 XX
 XX 09-DEC-1987; 87US-00130709.
 PR
 XX (MICR-) MICROPROBE CORP.
 PA
 XX Schwartz DE;
 PI
 XX WPI; 1989-192703/26.
 DR
 XX Vol. exclusion agents for enhancing in situ hybridisation - used for in
 PT vol. determ. of poly-nucleotide targets in cells by fixing cells to
 PT inert transparent solid support and hybridising, etc.
 XX
 PS Disclosure; Page 22; 28pp; English.
 XX
 XX The probe is complementary to Human Papillomavirus (HPV) mRNA. It detects
 CC bases 420-444 in gene E6 of HPV type 16. See also AAN93011-N93031.
 CC (Updated on 31-OCT-2002 to add missing OS field.)
 CC
 XX Sequence 24 BP; 5 A; 6 C; 6 G; 7 T; 0 U; 0 Other;
 SQ
 Query Match 100.0%; Score 21; DB 1; Length 24;
 Best Local Similarity 100.0%; Pred. No. 1;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CAGGACACAGTGGCTTTGAC 21
 Db |||||
 4 CAGGACACAGTGGCTTTGAC 24
 RESULT 6
 AAN91584
 ID AAN91584 standard; DNA; 24 BP.
 XX
 XX AAN91584;
 AC
 XX 27-AUG-2003 (revised)
 DT 17-JUL-1990 (first entry)
 XX
 XX Probe 6 for the detection of human papilloma virus (HPV) type 16 in in
 DE situ hybridisation assay.
 XX
 KM Human papilloma virus; type 16; in situ hybridisation assay;
 KM cellular smear; probe 6; benign wart; cervical cancer.
 XX
 OS Human papillomavirus.
 XX
 XX WO8902934-A.
 PN
 XX 06-APR-1989.
 PD
 XX 30-SEP-1988; 88WO-US003367.
 PF
 XX 02-OCT-1987; 87US-00103979.
 PR
 XX

PA (MICR-) MICROPROBE CORP.
 XX
 XX Schwartz DE; Adams TH;
 PI
 XX WPI; 1989-114406/15.
 DR
 XX Hybridisation test for human papilloma virus in cell smears - by reaction
 PT with long labelled probe specific for particular virus types, esp. for
 PT examining cervical smears.
 XX
 XX Disclosure; Page 19; 39pp; English.
 PS
 XX The patent is for a rapid in situ hybridisation assay for detecting and
 CC typing human papilloma virus (HPV) in non-frozen cellular smears fixed to
 CC a support in absence of aldehyde-based crosslinking reagents. The assay
 CC comprises: (1) combining nucleic acid in the sample with at least one
 CC detectable probe able to hybridise with 1 or more HPV types; and (2)
 CC detecting presence or absence of hybrid complexes. Opt. several probes
 CC are used, eg one for HPV types 6 and 11, associated with benign warts,
 CC and one for types 16, 18, 31, 33 and 35, associated with cervical cancer.
 CC The assay can differentiate between HPV types. It is esp. used as a
 CC secondary test. The probes can be synthesised or cloned. (Updated on 27-
 CC AUG-2003 to correct OS field.)
 CC
 XX Sequence 24 BP; 5 A; 6 C; 6 G; 7 T; 0 U; 0 Other;
 SQ
 Query Match 100.0%; Score 21; DB 1; Length 24;
 Best Local Similarity 100.0%; Pred. No. 1;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CAGGACACAGTGGCTTTGAC 21
 Db |||||
 4 CAGGACACAGTGGCTTTGAC 24
 RESULT 7
 AAV3991
 ID AAV3991 standard; DNA; 100 BP.
 XX
 XX AAV3991;
 AC
 XX 08-OCT-1998 (first entry)
 DT
 XX B6 oncogene from HPV 16 target sequence.
 DE
 XX Human papillomavirus; target; probe; detection; hybridisation;
 KM looped nucleic acid; diagnosis; pathogen; ss.
 KM
 XX Human papillomavirus.
 OS
 OS Synthetic.
 OS
 PN WO827225-A1.
 PD
 PD 25-JUN-1998.
 PF
 XX 16-DEC-1997; 97WO-GB003449.
 XX
 PR 16-DEC-1996; 96GB-00026074.
 XX
 PA (CYTO-) CYTOCELL LTD.
 XX
 XX Cardy DLM;
 PI
 XX WPI; 1998-362792/31.
 DR
 XX Detecting target nucleic acid by hybridisation with probe that generates
 PT looped-out region - then using second probe, hybridised to this region,
 PT to initiate nucleic acid synthesis, for disease diagnosis, detecting
 PT pathogens etc.
 XX
 XX Example 4; Page 19; 48pp; English.
 PS
 XX A method has been developed for detecting a target nucleic acid in a
 CC

XX WPI; 2002-566693/60.
XX
XX Novel cell having pre-trans-splicing molecules with target binding
PT domains that target binding of PTM to pre-mRNA, 3' or 5' splice region,
PT spacer region, nucleotide sequence to be trans-spliced to target-pre-
PT mRNA.
XX
XX Example; Fig 53; 229pp; English.
XX
XX The present invention describes a cell (I) comprising pre-trans-splicing
CC molecules (PTM) (II) which have one or more target binding domains (IIB)
CC that target binding of PTM to pre-mRNA, 3' splice region (IIB) that
CC includes branch point pyrimidine tract and 3' splice acceptor site, or 5'
CC splice site (IIC), spacer region (IID) that separates RNA splice site
CC from target binding domain, and nucleotide sequence to (IIE) be trans-
CC spliced to target-pre-mRNA. Optionally, the cell comprises (II) either
CC comprising: (A) (IIB) and (IIE); or (B) (IIC), (IID) and (IIE). The cell
CC may comprise a recombinant vector expressing (II). (I) has cytosol, the
CC immunosuppressive and antimicrobial activities, and can be used in gene
CC therapy. (II) comprising one or more (preferably two or more) (IIA) and
CC (IIB) (or (IIC)) and (IIE), or (II) comprising either (A) or (B)
CC (excluding (IIB)), is useful for producing a chimeric RNA molecule in a
CC cell which involves contacting a target pre-mRNA expressed in the cell
CC with (II) that is recognized by nuclear splicing components. The chimeric
CC RNA produced comprises sequences encoding a toxin or translatable
CC protein. The nucleotide sequence to be trans-spliced to target pre-mRNA
CC preferably comprises nucleotide sequences comprising exons 1-10 of cystic
CC fibrosis trans-membrane conductance regulator (CFTR). The chimeric RNA
CC molecule produced using (II) which either comprises (A) or (B) further
CC comprises a nucleotide sequence tag. (I) can be used for gene regulation,
CC gene repair and targeted cell death. (I) can be used for the treatment of
CC various diseases including genetic, infectious or autoimmune diseases and
CC proliferative disorders such as cancer and to regulate gene expression in
CC plants. ABQ73414 to ABQ73536 represent sequences used in the
CC exemplification of the present invention
XX
XX
SQ Sequence 193 BP; 46 A; 32 C; 22 G; 40 T; 0 U; 53 Other;
XX
XX
Query Match 100.0%; Score 21; DB 6; Length 193;
Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 CAGGACACAGTGGCTTTGAC 21
DB 92 CAGGACACAGTGGCTTTGAC 112
XX
RESULT 10
ADX26909/c
ID ADX26909 standard; DNA; 452 BP.
XX
XX ADX26909;
AC
XX 05-MAY-2005 (first entry)
XX
XX Human papillomavirus type 16 E6 delta151 mutant DNA.
DB
XX cell culture; immortalization; Gynecology and obstetrics; Andrology;
KM cancer; neoplasm; mutant; ds.
XX
XX Human papillomavirus type 16.
OS Synthetic.
XX
XX Key Location/Qualifiers
FH misc_difference 3..4
FT /*tag= a
FT /note= "Wild-type bases 4-25 have been removed"
FT misc_difference 449..450
FT /*tag= b
FT /note= "Wild-type bases 472-474 have been removed"
XX
XX JP2005046117-A.
XX
XX

XX
PD 24-FEB-2005.
XX
XX 31-JUL-2003; 2003JP-00283911.
PF
XX 31-JUL-2003; 2003JP-00283911.
XX
XX 31-JUL-2003; 2003JP-00283911.
XX
XX (UYKA-) UNIV KANAZAWA TLO YG.
PA (KOKU-) KOKURITSU GAN CENT SOCHO.
XX
XX WPI; 2005-175815/19.
DR
XX
XX The invention relates to a novel immortalized endometrial glandular
CC epithelial cell strain in which an exogenous immortalizing gene is
CC introduced and the properties of the cell line as non-cancer typical
CC glandular epithelial cells, are maintained. The cell line of the
CC invention may be useful for research of reproduction, particularly for
CC analyzing the adhesion of the fertilized egg during implantation,
CC analyzing the effect of steroid hormones, analyzing the mechanism of
CC multistep canceration of the inner glandular epithelial membrane by
CC genetic engineering of the cell line which is otherwise non-cancerous,
CC and analyzing the effect of carcinogenic substances on the cell line. The
CC cell line maintains the properties of normal somatic cells and the
CC properties of endometrial glandular epithelial cells. The current
CC sequence is that of the Human papillomavirus type 16 E6 delta151 mutant
CC DNA of the invention.
XX
XX
SQ Sequence 452 BP; 156 A; 70 C; 98 G; 128 T; 0 U; 0 Other;
XX
XX
Query Match 100.0%; Score 21; DB 14; Length 452;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 CAGGACACAGTGGCTTTGAC 21
DB 336 CAGGACACAGTGGCTTTGAC 316
XX
RESULT 11
ABQ76223/c
ID ABQ76223 standard; DNA; 456 BP.
XX
XX ABQ76223;
AC
XX 21-OCT-2002 (first entry)
XX
XX Human papillomavirus E6 tumour antigen derived DNA.
DE
XX Tumour antigen; human; vaccine; cellular immune response; immunogen; E6;
KM cancer; tumour; ds.
XX
XX Human papillomavirus.
OS
XX US6287569-B1.
PN
XX 11-SEP-2001.
XX
XX 06-APR-1998; 98US-00056105.
PF
XX 10-APR-1997; 97US-0043467P.
PR
XX (REGC) UNIV CALIFORNIA.
XX
XX Kippe TJ, Wu Y;
PI
XX WPI; 1998-583198/49.
XX
XX

XX Generating cellular immune response in patient to target protein -
PT comprises introducing vector with nucleotide sequence encoding immunogen
PT comprising protein processing signal into cell of patient.
XX
XX Disclosure; Col 91-92; 61pp; English.
XX This invention describes a novel method for generating a cellular immune
CC response in a patient to a target protein or its fragment. The method
CC involves introducing a vector containing a nucleotide sequence encoding a
CC chimeric immunogen comprising a protein processing signal and the target
CC protein or its fragment. The immunogen is produced by the cells and
CC processed so that the target protein or its fragment is presented to the
CC patient's immune system and a cellular immune response is initiated. The
CC method and vectors can be used as a form of vaccination and could be used
CC to generate a cellular immune response in patients to, e.g. cancerous
CC tumours. The cellular immune response is the predominant immune response
CC in the patient. This sequence represents a DNA fragment which encodes a
CC tumour antigen from human papillomavirus E6 described in the method of
CC the invention. Note: The information in this spec has been previously
CC disclosed in WO199845444 however this spec contained no sequence
CC information
XX
SQ Sequence 456 BP; 156 A; 70 C; 101 G; 129 T; 0 U; 0 Other;
Query Match 100.0%; Score 21; DB 2; Length 456;
Best Local Similarity 100.0%; Pred. No. 1.7; Mismatches 0; Gaps 0;
Matches 21; Conservative 0; Indels 0; Gaps 0;
QY 1 CAGGACACAGTGGCTTTGAC 21
DB 337 CAGGACACAGTGGCTTTGAC 317
RESULT 12
ADC37149/c
ID ADC37149 standard; DNA; 459 BP.
XX
XX ADC37149;
XX
XX 18-DEC-2003 (first entry)
XX
XX 459bp HPV16 subgenomic fragment DNA sequence.
XX
XX
XX immortalised non-tumorigenic; human Schwann; schwannoma cell line;
XX screening cancer chemotherapeutic; antineoplastic activity;
XX carcinogenicity; neuroprotective; neurodegeneration; SV40;
XX Simian Virus 40; adenovirus; human papilloma virus; E6; E7; cytostatic;
XX gene therapy; cancer; HPV16; ds.
XX
XX
XX Human papillomavirus type 16.
XX
XX WO2003073996-A2.
XX
XX 12-SEP-2003.
XX
XX 03-MAR-2003; 2003WO-US006314.
XX
XX 01-MAR-2002; 2002US-0361528P.
XX
XX (Hous-) HOUSE EAR INST.
XX
XX Hung G, Li X;
XX
XX WPI; 2003-722006/68.
XX
XX
XX Producing an immortalized non-tumorigenic human Schwann or schwannoma
PT cell line for treating cancer or neurodegenerative disorder comprises
PT selecting for immortalized cells that express the exogenous immortalizing
PT gene.
XX
XX Disclosure; SEQ ID NO 1; 29pp; English.
XX

CC The invention relates to a the production of an immortalised non-
CC tumorigenic human Schwann or schwannoma cell line comprising: providing
CC a cell culture of human Schwann or schwannoma cells, introducing a
CC polynucleotide comprising an exogenous immortalising gene into the cells,
CC and selecting for immortalised cells that express the exogenous
CC immortalising gene and retain phenotypic properties of Schwann or
CC schwannoma cells. The invention further comprises: determining the effect
CC of a pharmacological agent on human Schwann or schwannoma cells;
CC screening cancer chemotherapeutic and antineoplastic activity of an agent
CC ; testing carcinogenicity of an agent; screening the neuroprotective
CC activity of an agent; treating neurodegeneration in a patient; and a kit
CC for screening a pharmacological agent on schwannoma cells. In the method
CC of producing an immortalised non-tumorigenic human Schwann or schwannoma
CC cell line, the polynucleotide is a subgenomic fragment of a virus
CC consisting of SV40 (Simian Virus 40), adenovirus and human papilloma
CC virus, comprising the E6 and E7 genes of the human papilloma virus
CC consisting of types 18, 31, 33 or 35, but preferably 16. The immortalised
CC non-tumorigenic human Schwann or schwannoma cell line has cytostatic and
CC neuroprotective activity and be useful in treating disorders by gene
CC therapy. The method is useful for producing an immortalised non-
CC tumorigenic human Schwann or schwannoma cell line useful for screening
CC cancer chemotherapeutic and antineoplastic activity of an agent, testing
CC carcinogenicity of an agent and screening a neuroprotective activity of
CC an agent and for preparing a composition for diagnosing or treating
CC neurodegenerative disorders and cancer. This polynucleotide represents
CC the DNA of a 459bp HPV16 sequence, a subgenomic fragment used in the
CC production of the immortalised non-tumorigenic human Schwann or
CC schwannoma cell line of the invention.
XX
SQ Sequence 459 BP; 157 A; 71 C; 100 G; 131 T; 0 U; 0 Other;
Query Match 100.0%; Score 21; DB 10; Length 459;
Best Local Similarity 100.0%; Pred. No. 1.7; Mismatches 0; Gaps 0;
Matches 21; Conservative 0; Indels 0; Gaps 0;
QY 1 CAGGACACAGTGGCTTTGAC 21
DB 338 CAGGACACAGTGGCTTTGAC 318
RESULT 13
AD171963/c
ID AD171963 standard; DNA; 464 BP.
XX
XX
XX AD171963;
XX
XX 20-MAY-2004 (first entry)
XX
XX
XX Human ovarian cancer DNA marker #4705.
XX
XX
XX Human; ovarian cancer; ds; tumour; cytostatic; DNA marker.
XX
XX Homo sapiens.
XX
XX WO200170979-A2.
XX
XX 27-SEP-2001.
XX
XX 21-MAR-2001; 2001WO-US009126.
XX
XX 21-MAR-2000; 2000US-0191031P.
XX 25-MAY-2000; 2000US-0207124P.
XX 15-JUN-2000; 2000US-0211940P.
XX 07-JUL-2000; 2000US-0216820P.
XX 25-JUL-2000; 2000US-0220651P.
XX 21-DEC-2000; 2000US-0257672P.
XX
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
XX Lee J, Little J;
XX
XX WPI; 2001-611502/70.
XX

PT Novel isolated nucleic acid molecules (markers) overexpressed in ovarian
 PT cancer cells as compared to their normal non-cancerous ovarian cells are
 PT used to characterize stage, grade, histological type of ovarian cancer.
 XX
 XX Disclosure; SEQ ID NO 4705; 106pp; English.
 XX
 CC The invention relates to nucleic acid markers which are overexpressed in
 CC ovarian cancer cells as compared to their expression in normal (i.e. non-
 CC cancerous) ovarian cells. The invention also relates to polypeptides
 CC encoded by the markers, antibodies that selectively bind to the
 CC polypeptides, a method of inhibiting ovarian cancer in a patient at risk
 CC of developing ovarian cancer involving inhibiting expression of a gene
 CC corresponding to a marker of the invention and a method of treating a
 CC patient afflicted with ovarian cancer comprising providing to cells of
 CC the patient an antisense oligonucleotide complementary to a marker of the
 CC invention. The markers are useful for assessing if a patient is afflicted
 CC with ovarian cancer, which involves comparing the level of expression of
 CC a marker in a patient sample and a normal level of expression of the
 CC marker in a control non-ovarian cancer sample. A difference between the
 CC expression levels indicates ovarian cancer. The level of expression of a
 CC marker corresponds to a secreted protein or to a transcribed
 CC polynucleotide or its portion. The level of expression of the marker is
 CC assessed by detecting the presence in the sample, a protein or protein
 CC fragment corresponding to the marker. The presence of protein or protein
 CC fragment is detected using an antibody that specifically binds with the
 CC protein or protein fragment. Alternatively, the level of expression of
 CC the marker is assessed by detecting the presence of a transcribed
 CC polynucleotide which anneals with the marker or anneals with a portion of
 CC the polynucleotide comprising the marker, under stringent conditions. The
 CC marker is also used for monitoring the progression of ovarian cancer in a
 CC patient which involves detecting expression of the marker in a patient
 CC sample at a first point in time, repeating the method at a subsequent
 CC time and comparing the level of expression. The method is carried out
 CC using an ovarian tissue sample. A composition comprising a marker.
 CC polypeptide or antibody of the invention is used to treat ovarian cancer.
 CC This sequence represents a human ovarian cancer DNA marker of the
 CC invention. Note: The sequence data for this patent did not form part of
 CC the printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 464 BP; 139 A; 106 C; 117 G; 100 T; 0 U; 2 Other;
 QY
 DB 1 CAGGACACAGTGGCTTTGAC 21
 88 CAGGACACAGTGGCTTTGAC 68
 Query Match 100.0%; Score 21; DB 5; Length 464;
 Best Local Similarity 100.0%; Pred. No. 1.7;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

PR 15-JUN-2000; 2000US-0211940P.
 PR 07-JUL-2000; 2000US-0216820P.
 PR 25-JUL-2000; 2000US-0220661P.
 PR 21-DEC-2000; 2000US-0257672P.
 XX
 XX (MILL-) MILLENNium PREDICTIVE MEDICINE INC.
 XX
 XX Lee J, Lillie J;
 XX WPI; 2001-611502/70.
 DR
 XX
 XX Novel isolated nucleic acid molecules (markers) overexpressed in ovarian
 PT cancer cells as compared to their normal non-cancerous ovarian cells are
 PT used to characterize stage, grade, histological type of ovarian cancer.
 XX
 XX Disclosure; SEQ ID NO 11004; 106pp; English.
 PS
 CC The invention relates to nucleic acid markers which are overexpressed in
 CC ovarian cancer cells as compared to their expression in normal (i.e. non-
 CC cancerous) ovarian cells. The invention also relates to polypeptides
 CC encoded by the markers, antibodies that selectively bind to the
 CC polypeptides, a method of inhibiting ovarian cancer in a patient at risk
 CC of developing ovarian cancer involving inhibiting expression of a gene
 CC corresponding to a marker of the invention and a method of treating a
 CC patient afflicted with ovarian cancer comprising providing to cells of
 CC the patient an antisense oligonucleotide complementary to a marker of the
 CC invention. The markers are useful for assessing if a patient is afflicted
 CC with ovarian cancer, which involves comparing the level of expression of
 CC a marker in a patient sample and a normal level of expression of the
 CC marker in a control non-ovarian cancer sample. A difference between the
 CC expression levels indicates ovarian cancer. The level of expression of a
 CC marker corresponds to a secreted protein or to a transcribed
 CC polynucleotide or its portion. The level of expression of the marker is
 CC assessed by detecting the presence in the sample, a protein or protein
 CC fragment corresponding to the marker. The presence of protein or protein
 CC fragment is detected using an antibody that specifically binds with the
 CC protein or protein fragment. Alternatively, the level of expression of
 CC the marker is assessed by detecting the presence of a transcribed
 CC polynucleotide which anneals with the marker or anneals with a portion of
 CC the polynucleotide comprising the marker, under stringent conditions. The
 CC marker is also used for monitoring the progression of ovarian cancer in a
 CC patient which involves detecting expression of the marker in a patient
 CC sample at a first point in time, repeating the method at a subsequent
 CC time and comparing the level of expression. The method is carried out
 CC using an ovarian tissue sample. A composition comprising a marker.
 CC polypeptide or antibody of the invention is used to treat ovarian cancer.
 CC This sequence represents a human ovarian cancer DNA marker of the
 CC invention.
 XX
 SQ Sequence 464 BP; 139 A; 106 C; 117 G; 100 T; 0 U; 2 Other;
 QY
 DB 1 CAGGACACAGTGGCTTTGAC 21
 88 CAGGACACAGTGGCTTTGAC 68
 Query Match 100.0%; Score 21; DB 5; Length 464;
 Best Local Similarity 100.0%; Pred. No. 1.7;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

PI Milner AJ;
XX WPI; 2003-221850/21.
XX
XX
XX Selective post-transcriptional silencing of an exogenous viral gene (e.g.
PT human papilloma virus (HPV) B6), for treating e.g. cancer, comprises
PT using a small interfering RNA (siRNA) construct homologous to an mRNA of
PT the gene.
XX
XX Claim 11; Fig 11; 44pp; English.
XX
CC The invention relates to a novel method for selective post-translational
CC silencing in a mammalian cell of the expression of an exogenous gene of
CC viral origin. The method comprises introducing into the cell a small
CC interfering RNA (siRNA) construct that is homologous to a part of the
CC mRNA sequence of the gene. The method is useful for the selective post-
CC transcriptional silencing of an exogenous gene of viral origin (e.g. an
CC oncogene or human papilloma virus (HPV) gene) in a mammalian cell. The
CC method or the siRNA is particularly useful for treating cancer, human
CC cervical cancer, human immunodeficiency virus (HIV), smallpox, flu,
CC common cold, or a disease caused by a HPV (e.g. genital warts, cervical
CC cancer, penile cancer, malignant squamous cell carcinomas or verruca
CC vulgaris). An siRNA construct or vector is useful for use as a medicament
CC for the diseases mentioned. The polynucleotide sequence of the invention
CC can be used to treat disorders by gene therapy. This polynucleotide
CC sequence represents the DNA of a HPV siRNA sequence of the invention
XX
SQ Sequence 477 BP; 167 A; 75 C; 104 G; 131 T; 0 U; 0 Other;
XX
Query Match 100.0%; Score 21; DB 10; Length 477;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 CAGGACACAGTGGCTTTGAC 21
DB 358 CAGGACACAGTGGCTTTGAC 338
XX
RESULT 18
AD26908/c
ID ADX26908 standard; DNA; 477 BP.
AC ADX26908;
XX
XX 05-MAY-2005 (first entry)
XX
XX Human papillomavirus type 16 B6 wild-type gene.
XX
XX cell culture; immortalization; Gynecology and obstetrics; Andrology;
XX cancer; neoplasm; ds.
XX
XX Human papillomavirus type 16.
XX
XX JP2005046117-A.
XX
XX 24-FEB-2005.
XX
XX 31-JUL-2003; 2003JP-00283911.
XX
XX 31-JUL-2003; 2003JP-00283911.
XX
XX (UYKA-) UNIV KANAZAWA TLO YG.
XX (KOKU-) KOKURITSU GAN CENT SOCHO.
XX
XX WPI; 2005-175815/19.
XX
XX Immortalized endometrial glandular epithelial cell strain containing
XX exogenous immortalizing gene and maintains properties of non-cancer cell
XX and typical glandular epithelial cells, useful in field of research of
XX reproduction.
XX
XX Claim 4; SEQ ID NO 3; 22pp; Japanese.
XX

CC The invention relates to a novel immortalized endometrial glandular
CC epithelial cell strain in which an exogenous immortalizing gene is
CC introduced and the properties of the cell line as non-cancer typical
CC glandular epithelial cells, are maintained. The cell line of the
CC invention may be useful for research of reproduction, particularly for
CC analyzing the adhesion of the fertilized egg during implantation,
CC analyzing the effect of steroid hormones, analyzing the mechanism of
CC multistep cancerization of the inner glandular epithelial membrane by
CC genetic engineering of the cell line which is otherwise non-cancerous,
CC and analyzing the effect of carcinogenic substances on the cell line. The
CC cell line maintains the properties of normal somatic cells and the
CC properties of endometrial glandular epithelial cells. The current
CC sequence is that of the Human papillomavirus type 16 B6 wild-type gene of
CC the invention.
XX
SQ Sequence 477 BP; 167 A; 76 C; 103 G; 131 T; 0 U; 0 Other;
XX
Query Match 100.0%; Score 21; DB 14; Length 477;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 CAGGACACAGTGGCTTTGAC 21
DB 358 CAGGACACAGTGGCTTTGAC 338
XX
RESULT 19
AEA5113/c
ID AEA5113 standard; DNA; 477 BP.
AC AEA5113;
XX
XX 11-AUG-2005 (first entry)
XX
XX Human papillomavirus type 16 B6/B7 gene fragment. SEQ ID NO: 17.
XX
XX Delivery mechanism; gene therapy; cytostatic; vulnerary; virucide;
XX injury; infection; cancer; gastrointestinal disease;
XX gynecology and obstetrics; tumor; colorectal tumor;
XX uterine cervix tumor; squamous cell carcinoma; neoplasm; E7 gene; ds;
XX B6 gene.
XX
XX Human papillomavirus.
XX
XX WO2005051431-A1.
XX
XX 09-JUN-2005.
XX
XX 25-NOV-2004; 2004WO-GB004979.
XX
XX 25-NOV-2003; 2003GB-00027409.
XX
XX 05-MAR-2004; 2004US-0549919P.
XX
XX (MILN/) MILNER A J.
XX
XX Milner AJ;
XX
XX WPI; 2005-405310/41.
XX
XX Composition useful for delivering an agent into a cell comprises the
XX agent, a transfer agent and a solid or colloidal carrier medium.
XX
XX Disclosure; SEQ ID NO 17; 56pp; English.
XX
XX The present invention relates to a method and composition comprising a
XX transfer agent and a solid or colloidal carrier medium for delivering a
XX biological agents into cells. The invention is useful for the treatment
XX of cancer, tumor, carcinoma of cutaneous, squamous or cervical epithelia
XX and colorectal carcinoma, wounds, burns and scars. The invention is also
XX useful in gene therapy. The present sequence is the human papillomavirus
XX type 16 B6/B7 gene fragment. This sequence is useful in a method for
XX treating carcinomas.
XX

```

SQ Sequence 477 BP; 167 A; 75 C; 104 G; 131 T; 0 U; 0 Other;
Query Match      100.0%; Score 21; DB 14; Length 477;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      1 CAGGACACAGTGGCTTTTGAC 21
      |||||
Db      358 CAGGACACAGTGGCTTTTGAC 338

RESULT 20
AAT31834/c
ID AAT31834 standard; DNA; 519 BP.
XX
AC AAT31834;
XX
DT 27-AUG-2003 (revised)
DT 11-JAN-1997 (first entry)
XX
DE Human papilloma virus E6/E7 protein variant.
XX
KM Human papilloma virus; E6; E7; deletion mutant; HPV; immune response;
XX humoral immune response; cellular immune response; vaccine; ss.
XX
OS Human papillomavirus.
XX
FH Key Location/Qualifiers
FT CDS 1..519
FT FT /*tag= a
FT FT /product= "E6/E7 fusion protein (deletion mutant)."

```

```

FT FT misc_feature 154..156
FT FT /*tag= p
FT FT /transl_except= CAA encodes Ile
FT FT misc_feature 157..159
FT FT /*tag= q
FT FT /transl_except= AAG encodes Arg
FT FT misc_feature 160..162
FT FT /*tag= r
FT FT /transl_except= CCA encodes Cys
XX
XX WO9619496-A1.
XX
XX 27-JUN-1996.
XX
XX 20-DEC-1995; 95WO-AU000868.
XX
XX 20-DEC-1994; 94AU-00000157.
XX
XX (CSLC-) CSL LTD.
XX (UYQU) UNIV QUEBENSAND.
XX
XX Edwards SJ, Cox J, Webb EA, Frazer I;
XX
XX WPI; 1996-309518/31.
XX
XX P-PSDB; AAR97562.
XX
XX Vaccine variants of human papilloma virus antigens - contain variants of
XX E6 and/or E7 protein, pref. deletion mutants, and are used to treat or
XX prevent HPV infection.
XX
XX Example 3; Page 17; 37pp; English.
XX
XX A variant of the human papilloma virus (HPV) E6 or E7 protein which
XX elicits a humoral and/or cellular immune response against HPV can be used
XX in vaccines against HPV or to treat HPV infection. The variant is
XX preferably a deletion mutant comprising at least half, and preferably two
XX thirds of full length E6 or E7 protein starting from the N- or C-
XX terminal, or is a full length E6 moiety fused to a full length E7 moiety.
XX The variant optionally has a linkage moiety and a foreign protein or
XX peptide which facilitates the purification of, and enhances the
XX immunogenicity of, the fusion protein. This sequence encodes a fusion
XX between the C-terminal end of E6 and the N-terminal end of E7. The
XX protein is also a deletion mutant generated from the sequence described
XX in AAT31833. (Updated on 27-AUG-2003 to correct OS field.)
XX
XX SQ Sequence 519 BP; 175 A; 92 C; 113 G; 139 T; 0 U; 0 Other;
Query Match      100.0%; Score 21; DB 2; Length 519;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      1 CAGGACACAGTGGCTTTTGAC 21
      |||||
Db      172 CAGGACACAGTGGCTTTTGAC 152

RESULT 21
ADF31984/c
ID ADF31984 standard; DNA; 543 BP.
XX
XX ADF31984;
XX
XX 12-FEB-2004 (first entry)
XX
XX Human papillomavirus fusion gene encoding sequence.
XX
XX human papillomavirus; cervix cancer; ds.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 1..543
XX FT /*tag= a

```

XX CN1381583-A.
 XX
 PD 27-NOV-2002.
 XX
 PF 24-APR-2002; 2002CN-00117143.
 XX
 PR 24-APR-2002; 2002CN-00117143.
 XX
 PA (ONCO-) INST ONCOLOGY TUMOR HOSPITAL CHINESE ACA.
 XX
 PI Zhao Q;
 XX
 DR WPI; 2003-258260/26.
 DR P-PSDB; ADF31985.
 XX
 PT Human papillomavirus B6/E7 fusion gene and its efficient expression
 carrier and fusion protein vaccine.
 XX
 PS Clatm 3; SEQ ID NO 1; 16pp; Chinese.
 XX
 CC The present invention relates to human papillomavirus B6/E7 fusion gene,
 its preparing process, the process for configuring the efficient
 expression carrier containing the gene and resultant expression carrier,
 the fusion protein prepared from the gene, and the application of the
 fusion gene and expression protein to medical science and medicine to
 treat cervix cancer are disclosed. The present sequence represents the
 human papillomavirus fusion gene encoding sequence.
 CC
 SQ Sequence 543 BP; 185 A; 89 C; 118 G; 151 T; 0 U; 0 Other;

Query Match 100.0%; Score 21; DB 10; Length 543;
 Best Local Similarity 100.0%; Pred. No. 1.7;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAGGACACAGTGGCTTTGAC 21
 |||||
 DB 358 CAGGACACAGTGGCTTTGAC 338

RESULT 22
 AAQ75470/c
 ID AAQ75470 standard; DNA; 570 BP.
 XX
 AC AAQ75470;
 XX
 DT 16-OCT-2003 (revised)
 DT 25-MAR-2003 (revised)
 DT 28-JUN-1995 (first entry)
 XX
 DE HPV16 B6/E7 encoding region.
 XX
 KM HPV, HPV16, B6 protein; E7 protein; diagnosis; cervical dysplasia;
 KM cervix cancer; ds.
 XX
 OS Human papillomavirus; strain 16.
 XX
 FH Key Location/Qualifiers
 FT CDS 2..478
 FT /*tag= a
 FT /label= B6_encoding_region
 FT CDS 481..570
 FT /*tag= b
 FT /label= E7_encoding_region
 XX
 PN WO9426934-A2.
 PD 24-NOV-1994.
 XX
 PF 06-MAY-1994; 94MO-US005085.
 XX
 PR 06-MAY-1993; 93US-00058920.
 XX

PA (BAXT) BAXTER DIAGNOSTICS INC.
 XX
 PI Brown UT;
 XX
 DR WPI; 1995-006821/01.
 DR P-PSDB; AAR63865.
 XX
 PT Human papilloma virus detection assay - by amplification using self
 sustained sequence replication and hybridisation with a detector probe.
 XX
 PS Disclosure; Page 24-26; 79pp; English.
 XX

CC The sequences of the B6 and E7 polypeptide-encoding regions of human
 CC papillomavirus (HPV) 16 and 18 are given in AAQ75470-71 and the encoded
 CC proteins in AAR63865-66, respectively. Probes and primers based on these
 CC sequences were used for HPV infection diagnosis; expression of B6 and E7
 CC is diagnostic for cervical cancer or pre-malignant states. (Updated on
 CC 25-MAR-2003 to correct PN field.) (Updated on 16-OCT-2003 to standardise
 CC OS field)
 CC

SQ Sequence 570 BP; 199 A; 93 C; 119 G; 159 T; 0 U; 0 Other;

Query Match 100.0%; Score 21; DB 2; Length 570;
 Best Local Similarity 100.0%; Pred. No. 1.8;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAGGACACAGTGGCTTTGAC 21
 |||||
 DB 359 CAGGACACAGTGGCTTTGAC 339

RESULT 23
 ADL43498/c
 ID ADL43498 standard; DNA; 598 BP.
 XX

AC ADL43498;
 XX

DT 20-MAY-2004 (first entry)
 XX

DE Human ovarian cancer DNA marker #17388.
 XX

KM Human; ovarian cancer; ds; tumour; cytostatic; DNA marker.
 XX

OS Homo sapiens.
 XX

PN WO200170979-A2.
 XX

DT 27-SEP-2001.
 XX

PF 21-MAR-2001; 2001MO-US009126.
 XX

PR 21-MAR-2000; 2000US-0191031P.
 PR 25-MAY-2000; 2000US-0207124P.
 PR 15-JUN-2000; 2000US-0211940P.
 PR 07-JUL-2000; 2000US-0216820P.
 PR 25-JUL-2000; 2000US-0220661P.
 PR 21-DEC-2000; 2000US-0257672P.
 XX

PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
 XX

PI Lee J, Lillie J;
 XX

DR WPI; 2001-611502/70.
 XX

PT Novel isolated nucleic acid molecules (markers) overexpressed in ovarian
 cancer cells as compared to their normal non-cancerous ovarian cells are
 used to characterize stage, grade, histological type of ovarian cancer.
 XX
 PS Disclosure; SEQ ID NO 17388; 106pp; English.
 XX

CC The invention relates to nucleic acid markers which are overexpressed in
 CC ovarian cancer cells as compared to their expression in normal (i.e. non-
 CC cancerous) ovarian cells. The invention also relates to polypeptides

CC patient afflicted with ovarian cancer comprising providing to cells of
CC the patient an antisense oligonucleotide complementary to a marker of the
CC invention. The markers are useful for assessing if a patient is afflicted
CC with ovarian cancer, which involves comparing the level of expression of
CC a marker in a patient sample and a normal level of expression of the
CC marker in a control non-ovarian cancer sample. A difference between the
CC expression levels indicates ovarian cancer. The level of expression of a
CC marker corresponds to a secreted protein or to a transcribed
CC polynucleotide or its portion. The level of expression of the marker is
CC assessed by detecting the presence in the sample, a protein or protein
CC fragment corresponding to the marker. The presence of protein or protein
CC fragment is detected using an antibody that specifically binds with the
CC protein or protein fragment. Alternatively, the level of expression of
CC the marker is assessed by detecting the presence of a transcribed
CC polynucleotide which anneals with the marker or anneals with a portion of
CC the polynucleotide comprising the marker, under stringent conditions. The
CC marker is also used for monitoring the progression of ovarian cancer in a
CC patient which involves detecting expression of the marker in a patient
CC sample at a first point in time, repeating the method at a subsequent
CC time and comparing the level of expression. The method is carried out
CC using an ovarian tissue sample. A composition comprising a marker,
CC polypeptide or antibody of the invention is used to treat ovarian cancer.
CC This sequence represents a human ovarian cancer DNA marker of the
CC invention. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 743 BP; 192 A; 177 C; 159 G; 204 T; 0 U; 11 Other;

Query Match 100.0%; Score 21; DB 5; Length 743;

Best Local Similarity 100.0%; Pred. No. 1.8; Mismatches 0; Gaps 0;

QY 1 CAGGACACAGTGGCTTTGAC 21
DB 576 CAGGACACAGTGGCTTTGAC 596

RESULT 26
AAH72718
ID AAH72718 standard; cDNA; 745 BP.

XX AAH72718;

DT 19-SEP-2001 (first entry)

XX Human cervical cancer marker nucleic acid 3992.

KW Cervical cancer; cytostatic; pre-malignant condition; gene therapy; ss.

XX Homo sapiens.

PN MO200142467-A2.

XX 14-UTN-2001.

PF 08-DEC-2000; 2000MO-US033312.

PR 08-DEC-1999; 99US-0169681P.

PR 21-DEC-1999; 99US-0171350P.

PR 14-MAR-2000; 2000US-0189315P.

PR 12-MAY-2000; 2000US-0203791P.

PR 09-JUN-2000; 2000US-0210600P.

PR 21-JUL-2000; 2000US-0220114P.

XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

PI Schlegel R, Deeds J, Berger A, Zhao X;

DR WPI; 2001-375006/39.

XX New isolated nucleic acid for diagnosing and treating cervical cancer and
PT for assessing and detecting compounds for treating the cancer.

XX Claim 1; Page 812; 1051bp; English.

XX The invention relates to novel genes (AAH68727-AAH73383) associated with
CC cervical cancer with cytostatic activity. The nucleic acids and encoded
CC polypeptides are useful: to assess if a patient is afflicted with
CC cervical cancer or has a pre-malignant condition; to monitor the
CC progression of cervical cancer or a premalignant condition in a patient;
CC and to select and/or assess the efficacy of a compound or therapy for
CC inhibiting cervical cancer in a patient. The nucleic acids may also be
CC useful for gene therapy

XX Sequence 745 BP; 190 A; 185 C; 158 G; 208 T; 0 U; 4 Other;

Query Match 100.0%; Score 21; DB 4; Length 745;

Best Local Similarity 100.0%; Pred. No. 1.8; Mismatches 0; Gaps 0;

QY 1 CAGGACACAGTGGCTTTGAC 21
DB 573 CAGGACACAGTGGCTTTGAC 593

RESULT 27
ADO44067/c
ID ADO44067 standard; DNA; 747 BP.

XX ADO44067;

DT 15-JUL-2004 (first entry)

XX Nucleotide sequence of an E7B6 fusion protein.

KW B6 protein; E7 protein; fusion protein; HPV6; HPV-associated cancer;
KW cervical cancer; immune response; lower gastrointestinal tract cancer;
KW anal cancer; reproductive system cancer; penile cancer; vulvar cancer;
KW gene; ss.

XX Human papillomavirus type 16.

OS Synthetic.

XX Key Location/Qualifiers

FT CDS 1..747

FT /tag= a

FT /product= "E7B6 fusion protein"

PN MO2004030636-A2.

XX 15-APR-2004.

PF 02-OCT-2003; 2003MO-US031726.

PR 03-OCT-2002; 2002US-0415929P.

PA (AMHP) WYETH HOLDINGS CORP.

PI Smith L, Casaccia MC;

DR WPI; 2004-316328/29.

DR P-PDB; ADO44066.

XX New polypeptide comprising human papillomavirus B6 and E7 polypeptides,
PT useful for treating or preventing human papillomavirus (HPV)-associated
PT cancers, e.g. cervical cancer.

XX Example 1; Page 73; 101bp; English.

XX The present sequence encodes an E7B6 fusion protein, comprising wild type
CC E7 and B6 polypeptides from human papillomavirus type 16 (HPV16). The
CC specification describes human papillomavirus B6 and E7 polypeptides,
CC where the E7 polypeptide has mutations at any one or more of the amino
CC acids corresponding to amino acids 24, 26 or 91 of the sequence given in
CC ADO44073 and the B6 polypeptide has no mutations or has mutations at any

CC one or more of the amino acids corresponding to amino acids 63 or 106 of
CC there sequence given in ADO44072. The polypeptides of the invention are
CC useful for treating or preventing human papillomavirus (HPV)-associated
CC cancers, such as cervical cancer. The fusion proteins and nucleic acids
CC encoding the fusion proteins are useful for generating immune responses
CC against HPV. They are also useful for treating lower gastrointestinal
CC tract cancers, e.g. anal cancer, and other cancers of the reproductive
CC system, including penile and vulvar cancer.
XX
SQ Sequence 747 BP; 247 A; 132 C; 166 G; 202 T; 0 U; 0 Other;
Query Match 100.0%; Score 21; DB 12; Length 747;
Best Local Similarity 100.0%; Pred. No. 1.8;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CAGGACACAGTGGCTTTTGAC 21
DB 628 CAGGACACAGTGGCTTTTGAC 608
RESULT 28
ADO44065/C
XX ADO44065 standard; DNA; 747 BP.
XX
AC ADO44065;
XX
XX 15-JUL-2004 (first entry)
XX
DE Nucleotide sequence of a fusion protein designated E6E7Pentm.
XX
XX B6 protein; E7 protein; fusion protein; HPV16; HPV-associated cancer;
XX cervical cancer; immune response; lower gastrointestinal tract cancer;
XX anal cancer; reproductive system cancer; penile cancer; vulvar cancer;
XX gene; ss.
XX
XX Human papillomavirus type 16.
XX Synthetic.
OS
XX
FH Key Location/Qualifiers
FT CDS 1..747
FT /tag= a
FT /product= "E6E7Pentm fusion protein"
XX
XX MO2004030636-A2.
XX
XX 15-APR-2004.
XX
XX 02-OCT-2003; 2003MO-US031726.
XX
XX 03-OCT-2002; 2002US-0415929P.
XX
XX (AMHP) WYETH HOLDINGS CORP.
XX
XX Smith L, Cassetti MC;
XX
XX WPI: 2004-316328/29.
XX P-PSDB; ADO44064.
XX
XX New polypeptide comprising human papillomavirus E6 and E7 polypeptides,
XX useful for treating or preventing human papillomavirus (HPV)-associated
XX cancers, e.g. cervical cancer.
XX
XX Claim 24, Page 71; 101pp; English.
XX
XX The present sequence encodes a fusion protein, comprising E6 and E7
XX polypeptides from human papillomavirus type 16 (HPV16). The fusion
XX protein is designated E6E7Pentm, and comprises an E6 amino terminus
XX (where residues 63 and 106 have been replaced with glycine) and an E7
XX carboxy terminus (where residues 24, 26 and 91 have been replaced with
XX glycine). E6E7Pentm is representative of fusion proteins of the
XX invention. The specification describes human papillomavirus E6 and E7
XX polypeptides, where the E7 polypeptide has mutations at any one or more
XX of the amino acids corresponding to amino acids 24, 26 or 91 of the

CC sequence given in ADO44073 and the E6 polypeptide has no mutations or has
CC mutations at any one or more of the amino acids corresponding to amino
CC acids 63 or 106 of there sequence given in ADO44072. The polypeptides of
CC the invention are useful for treating or preventing human papillomavirus
CC (HPV)-associated cancers, such as cervical cancer. The fusion proteins
CC and nucleic acids encoding the fusion proteins are useful for generating
CC immune responses against HPV. They are also useful for treating lower
CC gastrointestinal tract cancers, e.g. anal cancer, and other cancers of
CC the reproductive system, including penile and vulvar cancer.
XX
SQ Sequence 747 BP; 246 A; 132 C; 171 G; 198 T; 0 U; 0 Other;
Query Match 100.0%; Score 21; DB 12; Length 747;
Best Local Similarity 100.0%; Pred. No. 1.8;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CAGGACACAGTGGCTTTTGAC 21
DB 337 CAGGACACAGTGGCTTTTGAC 317
RESULT 29
ADO44061/C
XX ADO44061 standard; DNA; 747 BP.
XX
AC ADO44061;
XX
XX 15-JUL-2004 (first entry)
XX
XX Nucleotide sequence of an E6E7 fusion protein.
XX
XX B6 protein; E7 protein; fusion protein; HPV16; HPV-associated cancer;
XX cervical cancer; immune response; lower gastrointestinal tract cancer;
XX anal cancer; reproductive system cancer; penile cancer; vulvar cancer;
XX gene; ss.
XX
XX Human papillomavirus type 16.
XX Synthetic.
OS
XX
FH Key Location/Qualifiers
FT CDS 1..747
FT /tag= a
FT /product= "E6E7 fusion protein"
XX
XX MO2004030636-A2.
XX
XX 15-APR-2004.
XX
XX 02-OCT-2003; 2003MO-US031726.
XX
XX 03-OCT-2002; 2002US-0415929P.
XX
XX (AMHP) WYETH HOLDINGS CORP.
XX
XX Smith L, Cassetti MC;
XX
XX WPI: 2004-316328/29.
XX P-PSDB; ADO44060.
XX
XX New polypeptide comprising human papillomavirus E6 and E7 polypeptides,
XX useful for treating or preventing human papillomavirus (HPV)-associated
XX cancers, e.g. cervical cancer.
XX
XX Example 1; Page 68; 101pp; English.
XX
XX The present sequence encodes an E6E7 fusion protein, comprising wild type
XX E6 and E7 polypeptides from human papillomavirus type 16 (HPV16). The
XX specification describes human papillomavirus E6 and E7 polypeptides,
XX where the E7 polypeptide has mutations at any one or more of the amino
XX acids corresponding to amino acids 24, 26 or 91 of the sequence given in
XX ADO44073 and the E6 polypeptide has no mutations or has mutations at any
XX one or more of the amino acids corresponding to amino acids 63 or 106 of
XX there sequence given in ADO44072. The polypeptides of the invention are

CC one or more of the amino acids corresponding to amino acids 63 or 106 of
 CC the sequence given in ADO44072. The polypeptides of the invention are
 CC useful for treating or preventing human papillomavirus (HPV)-associated
 CC cancers, such as cervical cancer. The fusion proteins and nucleic acids
 CC encoding the fusion proteins are useful for generating immune responses
 CC against HPV. They are also useful for treating lower gastrointestinal
 CC tract cancers, e.g. anal cancer, and other cancers of the reproductive
 CC system, including penile and vulvar cancer.

XX Sequence 747 BP; 246 A; 132 C; 170 G; 199 T; 0 U; 0 Other;

SO Query Match 100.0%; Score 21; DB 12; Length 747;

Best Local Similarity 100.0%; Pred. No. 1.8;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAGGACACAGTGGCTTTTGAC 21
 |||||
 DB 628 CAGGACACAGTGGCTTTTGAC 608

RESULT 32

ADO44071/c
 ID ADO44071 standard; DNA; 747 BP.

XX ADO44071;

XX 15-UTR-2004 (first entry)

DE Nucleotide sequence of a fusion protein designated E7B6Pentm.

XX E6 protein; E7 protein; fusion protein; HPV16; HPV-associated cancer;

KW cervical cancer; immune response; lower gastrointestinal tract cancer;

KM anal cancer; reproductive system cancer; penile cancer; vulvar cancer;

KW gene; ss.

XX Human papillomavirus type 16.

OS Synthetic.

XX Key Location/Qualifiers

FT CDS 1..747

FT /product= "E7B6Pentm fusion protein"

XX WO2004030636-A2.

XX 15-APR-2004.

XX 02-OCT-2003; 2003WO-US031726.

XX 03-OCT-2002; 2002US-0415929P.

XX (AMHP) WYETH HOLDINGS CORP.

XX Smith L, Cassetti MC;

XX WPI; 2004-316328/29.

XX P-PDB; ADO44070.

XX New polypeptide comprising human papillomavirus E6 and E7 polypeptides,

XX useful for treating or preventing human papillomavirus (HPV)-associated

XX cancers, e.g. cervical cancer.

XX Claim 24; Page 76; 101pp; English.

XX The present sequence encodes a fusion protein, comprising E7 and E6

XX polypeptides from human papillomavirus type 16 (HPV16). The fusion

XX protein is designated E7B6Pentm, and comprises an E7 amino terminus

XX (where residues 24, 26 and 91 have been replaced with glycine) and an E6

XX carboxy terminus (where residues 63 and 106 have been replaced with

XX glycine). E7B6Pentm is representative of fusion proteins of the

XX invention. The specification describes human papillomavirus E6 and E7

XX polypeptides, where the E7 polypeptide has mutations at any one or more

XX of the amino acids corresponding to amino acids 24, 26 or 91 of the

CC sequence given in ADO44073 and the E6 polypeptide has no mutations or has
 CC mutations at any one or more of the amino acids corresponding to amino
 CC acids 63 or 106 of the sequence given in ADO44072. The polypeptides of
 CC the invention are useful for treating or preventing human papillomavirus
 CC (HPV)-associated cancers, such as cervical cancer. The fusion proteins
 CC and nucleic acids encoding the fusion proteins are useful for generating
 CC immune responses against HPV. They are also useful for treating lower
 CC gastrointestinal tract cancers, e.g. anal cancer, and other cancers of
 CC the reproductive system, including penile and vulvar cancer.

XX Sequence 747 BP; 246 A; 132 C; 171 G; 198 T; 0 U; 0 Other;

SO Query Match 100.0%; Score 21; DB 12; Length 747;

Best Local Similarity 100.0%; Pred. No. 1.8;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAGGACACAGTGGCTTTTGAC 21
 |||||
 DB 628 CAGGACACAGTGGCTTTTGAC 608

RESULT 33

ADR47004/c
 ID ADR47004 standard; DNA; 768 BP.

XX ADR47004;

XX 18-NOV-2004 (first entry)

DE Human papillomavirus type 16 E7 gene for Dengue virus vaccine.

XX de; gene; cytosolic; virucide; dengue virus; recombinant replicon;

KW deletion; prem protein; C protein; NS1 protein signal; vaccine;

KM cervical cancer; viral disease; antigen; dendritic cell; immune response;

KW human papillomavirus.

XX Human papillomavirus type 16.

OS Synthetic.

XX Key Location/Qualifiers

FT CDS 1..768

FT /product= "HPV-16 E7 protein"

XX WO2004072274-A1.

XX 26-AUG-2004.

XX 30-JAN-2004; 2004WO-CN000088.

XX 30-JAN-2003; 2003CN-00115272.

XX 30-JAN-2003; 2003CN-00115273.

XX (SHAN-) SHANGHAI TENGGEN BIOMEDICAL CO LTD.

XX (TENG-) TENGGEN BIOMEDICAL CO.

XX (BEIT-) BEIJING ORIENTAL TENGGEN TECHNOLOGY DEV C.

XX Pang X;

XX WPI; 2004-625870/60.

XX P-PDB; ADR47005.

XX GENBANK; AF486352.

XX A dengue virus recombinant replicon has a deletion of the complete coding

XX e.g. the non-coding region in the whole of the 5'-end, the coding region

XX of the front 20 amino acids in the C protein, and the coding region of

XX NS1 protein signal; coding regions of all non-structural proteins. The

obtained vaccines are useful in producing preventives or/and remedies for cancer like cervical cancer and viral diseases. Such vaccines can efficiently express antigen in infected cells, which is because dengue virus can infect dendritic cells, and can effectively present antigen to provide immunity effect. Different types of dengue virus can be used to repeatedly produce efficient immune response thereby strengthening the body's immune system against the pathogen that contains such antigen. Human papillomavirus (HPV) vaccines were prepared by using a gene-expressing system using of the full-length dengue virus cDNA clone (pBS/FMD2). The recombinant virus vectors were transfected into baby hamster kidney (BHK) cells to enable the screening of BHK-21 Tet-off cell lines. This sequence corresponds to the HPV type 16 E7 gene used as the gene of interest in the recombinant replicon of the invention.

SQ Sequence 768 BP; 255 A; 139 C; 172 G; 202 T; 0 U; 0 Other;

Query Match	100.0%	Score 21	DB 13	Length 768
Best Local Similarity	100.0%	Pred. NC	1.9	
Matches 21	Conservative 0	Mismatches 0	Indels 0	Gaps 0

OY		1	CAGGACACAGTGGCTTTTGAC	21
Db		652	CAGGACACAGTGGCTTTTGAC	632

RESULT 34
AEF40156/c
ID AEF40156 standard; DNA; 768 BP.

AC AEF40156;

DT 23-MAR-2006 (first entry)

DE Human papillomavirus 16 (HPV-16) E7-E6 oncogene.

KM Vaccine; virus-like particle; replicon; therapeutic; cancer; cytostatic;
KM neoplasm; viral infection; virucide; infection; oncogene;
KM coding sequence; ds.

OS Human papillomavirus type 16.

FT	Key	Location/Qualifiers
FT	CDS	1..768
FT		/*tag= a
FT		/partial
FT		/product= "Human papillomavirus 16 (HPV-16) E7-E6 oncoprotein"
FT		/note= "No stop codon"
FT		

PN US2006018928-A1.

PD 26-JAN-2006

PF 29-JUL-2005; 2005US-00192923.

PR 30-JAN-2003; 2003CN-00115272.

PR 30-JAN-2004; 2004WO-00072274.

PA (PANG/) PANG X.

PI Pang X;

DR WPI; 2006-109169/11.

DR GENBANK; AF486352, AF469197, AF472508.

PT New recombinant DEN replicons with a c

PT infection.

PS Example 2; SEQ ID NO 1; 24pp; English.

CC The present invention provides a virus-like particle (VLP) vaccine which
CC contains dengue virus (DENV) recombinant replicon as its core. The DEN
CC replicon contains exogenous nucleotide sequences such as human
CC papillomavirus (HPV) antigen proteins, immune regulators or combination
CC of HPV antigen and immune regulators. The invention is useful for
CC producing a drug for the prophylaxis and treatment of cancer or viral
CC infection. The present sequence is a human papillomavirus oncogene.

SQ Sequence 768 BP; 255 A; 139 C; 172 G; 202 T; 0 U; 0 Other;

Query Match	100.0%	Score 21	DB 15	Length 768
Best Local Similarity	100.0%	Pred. NC	1.9	
Matches 21, Conservative	0	Mismatches	0	Gaps 0

OY 1 CAGACACAAGTGGCTTTGAC 21
| | | | | | | | | |
Db 652 CAGACACAAGTGGCTTTGAC 632

RESULT 35
AAT14663/c
ID AAT14663 standard; DNA; 776 BP.

AC AAT14663;

DT 27-AUG-2003 (revised)

DT 10-OCT-1996 (first entry)

DE E6/E7 region of Human Papillomavirus (HPV) type 16 DNA.

KW Human papilloma virus; HPV; detection; cervical

22

[illegible]

FT	primer_bind	1.	.24
FT	primer_bind	1.	.24

```

PT misc_binding /note= "primer B113 binding site."
PT 30.55
PT /*tag= b
PT /note= "primer H16-58 binding site."
PT primer_bind 37.57
PT C
PT /*tag=
PT /note= "primer B84 binding site."

```

PN US5506105-A.

PD 09-APR-1996.

PF 22-MAR-1994; 94US-00216233.

```

XX 10-DEC-1991; 91US-00808456.
XX (DADE-) DADE INT INC.
XX Haydock PV;
XX WPI, 1996-200273/20.
XX Detection of low copy number intracellular markers - by 3SR amplification
XX of target RNA in fixed cells then hybridisation with labelled probe, for
XX detecting human papilloma virus in cervical cells.
XX
XX Example 2; Fig 3; 21pp; English.
XX
XX An in situ hybridisation assay for detecting an intracellular marker of
XX low copy number in cells comprises fixing the cells to a support using
XX paraformaldehyde; treating the cells with a protease to permeabilise them
XX without altering morphology; adding amplification reagents; incubating
XX the cells at below fifty degrees celsius to perform amplification by self
XX -sustained sequence replication; adding a labelled probe complementary to
XX the region between the primers; washing cells to remove unhybridised
XX probe and then detecting the labelled probe. The B6/E7 region of human
XX papilloma virus (HPV) 16 is used especially to detect mRNA being
XX transcribed from this region which becomes active in transformed cells.
XX The method can be used for the early diagnosis of cervical cancer.
XX Primers used to amplify fragments of the B6/E7 region are described in
XX AAT14664-T14674. (Updated on 25-MAR-2003 to correct PF field.) (Updated
XX on 27-AUG-2003 to correct OS field.)
XX
XX Sequence 776 BP; 260 A; 139 C; 169 G; 208 T; 0 U; 0 Other;
XX
XX Query Match 100.0%; Score 21; DB 2; Length 776;
XX Best Local Similarity 100.0%; Pred. No. 1.9;
XX Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 CAGGACACAGTGGCTTTTGAC 21
XX |||||||||||||||||||
XX 358 CAGGACACAGTGGCTTTTGAC 338
XX
XX RESULT 36
XX AAX89756/C
XX ID AAX89756 standard; DNA; 779 BP.
XX
XX AAX89756;
XX
XX 27-AUG-2003 (revised)
XX 12-OCT-1999 (first entry)
XX
XX Probe sequence for HPV 16 B6/E7.
XX
XX human papilloma virus; infection; gene expression; probe; detection;
XX assay; cancer; virus; HPV; ss.
XX
XX Synthetic.
XX Human papillomavirus.
XX
XX WO9292890-A2.
XX
XX 17-JUN-1999.
XX
XX 11-DEC-1998; 98WO-US026447.
XX
XX 12-DEC-1997; 97US-0069426P.
XX 05-JAN-1998; 98US-0070486P.
XX 17-APR-1998; 98US-0082167P.
XX
XX (DIGE-) DIGENE CORP.
XX
XX Lorincz AT;
XX
XX WPI, 1999-443850/37.

```

```

XX New method for assessing Human Papilloma Virus (HPV) infection by
XX comparison of gene expression levels.
XX
XX Disclosure; Fig 5; 35pp; English.
XX
XX This nucleotide probe is specific for the HPV16 Human Papilloma Virus
XX (HPV) gene. The degree of HPV infection can be assessed, by measuring the
XX levels of expression of genes involved in the diseased state, and
XX comparing the expression to each other or to reference genes. (Updated on
XX 27-AUG-2003 to correct OS field.)
XX
XX Sequence 779 BP; 259 A; 139 C; 170 G; 211 T; 0 U; 0 Other;
XX
XX Query Match 100.0%; Score 21; DB 2; Length 779;
XX Best Local Similarity 100.0%; Pred. No. 1.9;
XX Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 CAGGACACAGTGGCTTTTGAC 21
XX |||||||||||||||||||
XX 365 CAGGACACAGTGGCTTTTGAC 345
XX
XX RESULT 37
XX AAQ29389/C
XX ID AAQ29389 standard; DNA; 790 BP.
XX
XX AAQ29389;
XX
XX 24-OCT-2003 (revised)
XX 25-MAR-2003 (revised)
XX 09-MAR-1993 (first entry)
XX
XX DNA encoding HPV 16 B6/E7 proteins obtd. by PCR.
XX
XX Virus vector; vaccinia virus; papillomavirus; HPV; human; amplification;
XX immunotherapeutic; ss.
XX
XX Human papillomavirus type 16.
XX
XX Location/Qualifiers
XX Key 1..789
XX CDS /*tag= a
XX /note= "reading frame 1"
XX 2..790
XX /*tag= b
XX /note= "reading frame 2"
XX 3..788
XX /*tag= d
XX /note= "reading frame 3"
XX 6..482
XX /*tag= e
XX /note= "third reading frame encoding HPV 16 B6"
XX 485..781
XX /*tag= c
XX /note= "second reading frame encoding HPV 16 E7"
XX
XX WO9216636-A1.
XX
XX 01-OCT-1992.
XX
XX 10-MAR-1992; 92WO-GB000424.
XX
XX 14-MAR-1991; 91GB-00005383.
XX
XX (IMMU ) IMMUNOLOGY LTD.
XX
XX Bourneill MEG, Inglis SC, Munro AJ;
XX
XX WPI, 1992-349219/42.
XX P-PSDB; AAR27723, AAR27724, AAR27725.
XX
XX Recombinant virus vectors encoding human papillomavirus proteins - for

```

PT treating and vaccinating against HPV infections and conditions caused by
PT them, such as cervical cancer.
XX
PS Disclosure; Fig 1a; 83pp; English.
XX
CC The fragment of DNA contg. the HPV-16 E6/E7 coding region was prep'd. by
CC PCR from plasmid BBR322/HPV16 (Durst et al., PNAS, 80: 3812 (1983)) using
CC oligonucleotides S05 and S06. The DNA prod. has a site for NcoI at the
CC beginning of the E6 gene and a SmaI site immediately downstream of the
CC termination codon for E7. The E6 and E7 ORFs are fused together to form a
CC single continuous ORF via site directed mutagenesis and the immortalising
CC potential of E7 is removed by altering two key codons of the HPV E7
CC sequence. The single ORF of HPV-16 E6/E7 may be inserted into vaccinia
CC virus DNA at neutral sites (pref. by inserting two sets of the DNA in
CC opposite orientations to overcome the problem of intertypic
CC recombination) to make a recombinant virus vector for use
CC immunotherapeutically to activate cells of the immune system against HPV.
CC See also AAG29385-400 and AAG29450-69. (Updated on 25-MAR-2003 to correct
CC PN field.) (Updated on 24-OCT-2003 to standardise OS field)
XX
SQ Sequence 790 BP; 262 A; 144 C; 175 G; 209 T; 0 U; 0 Other;
QY
Query Match 100.0%; Score 21; DB 2; Length 790;
Best Local Similarity 100.0%; Pred. No. 1.9;
Matches 21; Conservative 0; Mismatches 0; Gaps 0;
Indels 0;
DB 1 CAGACACAGTGGCTTTGAC 21
363 CAGACACAGTGGCTTTGAC 343
RESULT 38
AAT31833/c
ID AAT31833 standard; DNA; 801 BP.
XX
AC AAT31833;
XX
DT 27-AUG-2003 (revised)
DT 10-JAN-1997 (first entry)
XX
XX Human papilloma virus E6/E7 protein variant.
XX
DE Human papilloma virus E6/E7 protein variant.
XX
KW Human papilloma virus; E6; E7; deletion mutant; HPV; immune response;
KW humoral immune response; cellular immune response; vaccine; ss.
XX
OS Human papillomavirus.
XX
FH Key Location/Qualifiers
FT CDS 1..801
FT /tag= a
FT /product= "E6/E7 fusion protein."
XX
PN MO9619496-A1.
XX
PD 27-JUN-1996.
XX
PF 20-DEC-1995; 95MO-AU000868.
XX
PR 20-DEC-1994; 94AU-00000157.
XX
PA (CSLC-) CSL LTD.
PA (UYOU) UNIV QUEENSLAND.
XX
PI Edwards SJ, Cox J, Webb BA, Frazer I;
XX
DR WPI; 1996-309518/31.
DR P-PSDB; AAR97561.
XX
XX Vaccine variants of human papilloma virus antigens - contain variants of
PT E6 and/or E7 protein, pref. deletion mutants, and are used to treat or
PT prevent HPV infection.
XX
PS Example 1; Page 15-16; 37pp; English.

XX
CC A variant of the human papilloma virus (HPV) E6 or E7 protein which
CC elicits a humoral and/or cellular immune response against HPV can be used
CC in vaccines against HPV or to treat HPV infection. The variant is
CC preferably a deletion mutant comprising at least half, and preferably two
CC -thirds of full length E6 or E7 protein starting from the N- or C-
CC terminal, or is a full length E6 moiety fused to a full length E7 moiety.
CC The variant optionally has a linkage moiety and a foreign protein or
CC peptide which facilitates the purification of, and enhances the
CC immunogenicity of, the fusion protein. This sequence encodes a full
CC length E6/E7 fusion protein. (Updated on 27-AUG-2003 to correct OS
CC field.)
XX
SQ Sequence 801 BP; 266 A; 150 C; 174 G; 211 T; 0 U; 0 Other;
QY
Query Match 100.0%; Score 21; DB 2; Length 801;
Best Local Similarity 100.0%; Pred. No. 1.9;
Matches 21; Conservative 0; Mismatches 0; Gaps 0;
Indels 0;
DB 1 CAGACACAGTGGCTTTGAC 21
358 CAGACACAGTGGCTTTGAC 338
RESULT 39
AAK8792/c
ID AAK8792 standard; DNA; 822 BP.
XX
AC AAK8792;
XX
DT 06-SEP-1999 (first entry)
XX
DE HPV fusion protein DL/3-E6-His/HPV16 DNA.
XX
XX Fusion protein; E6 protein; E7 protein; E6/E7; immunomodulator; tumour;
KW immunological fusion partner; CpG oligonucleotide; immune response;
KW HPV antigen; prevention; treatment; ss.
XX
OS Synthetic.
XX
DE Human papillomavirus.
XX
PN WO933868-A2.
XX
PD 08-JUL-1999.
XX
PF 18-DEC-1998; 98MO-EP008563.
XX
PR 24-DEC-1997; 97GB-00027262.
XX
PA (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
XX
PI Dalemans WLJ, Gerard CMG;
XX
DR WPI; 1999-405485/34.
DR P-PSDB; AAY25376.
XX
PT Composition comprising an E6, E7 or E6/E7 fusion protein from HPV to
PT induce immune response to HPV.
XX
PS Example II; Page 48; 62pp; English.
XX
XX AAK8791-X78801 represent nucleic acid sequences which encode novel
CC constructs comprising an E6 or E7 protein or E6/E7 fusion protein from
CC HPV (represented in AAY25375-Y25386). These constructs are optionally
CC linked to an immunological fusion partner and an immunomodulatory CpG
CC oligonucleotide. The products of the invention can be used to induce an
CC immune response in a patient to an HPV antigen. They can also be used for
CC preventing or treating HPV induced tumours
XX
SQ Sequence 822 BP; 277 A; 147 C; 168 G; 230 T; 0 U; 0 Other;
QY
Query Match 100.0%; Score 21; DB 2; Length 822;
Best Local Similarity 100.0%; Pred. No. 1.9;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAGGACACAGTGGCTTTGAC 21
 |||||
 Db 676 CAGGACACAGTGGCTTTGAC 656

RESULT 40
 AAX29781/C
 ID AAX29781 standard; DNA; 822 BP.
 XX
 AC AAX29781;
 XX
 DT 17-OCT-2003 (revised)
 XX 22-JUN-1999 (first entry)
 XX
 DE Prot.D1/3-B6-H1s/HPV16 coding sequence.
 XX
 KM Chimeric; B6; E7; fusion protein; protein D; vaccine; immunotherapy;
 XX tumour; lesion; benign; malignant; virus; infection; ss.
 OS Human papillomavirus.
 OS Haemophilus influenzae.
 OS Chimeric.
 XX
 PN MO9910375-A2.
 XX
 PD 04-MAR-1999.
 XX
 PF 17-AUG-1998; 98MO-EP005285.
 XX
 PR 22-AUG-1997; 97GB-00017953.
 XX
 PA (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
 XX
 PI Bruck C, Cabezon Silva T, Delisse AEF, Gerard CMG;
 XX Lombardo-Bencheikh A;
 XX WPI; 1999-190587/16.
 DR
 XX
 PT Human Papilloma Virus (HPV) fusion proteins - useful in vaccines for
 XX treatment or prophylaxis of HPV induced lesions.
 XX
 PS Disclosure; Fig 3; 95pp; English.
 XX
 CC This sequence represents the coding region for a chimeric B6 or E7
 CC protein or B6/E7 fusion protein from Human papillomavirus (HPV) linked to
 CC an immunological fusion partner, in this case, a fragment of the
 CC Haemophilus influenzae B protein D. The sequence also contains a
 CC histidine tag at the C-terminus of the encoded protein. The protein can
 CC be used in a vaccine, for immuno-therapeutically treating HPV induced
 CC tumour lesions (benign or malignant) and preventing HPV viral infection.
 CC (Updated on 17-OCT-2003 to standardise OS field)
 CC
 SQ Sequence 822 BP; 277 A; 147 C; 168 G; 230 T; 0 U; 0 Other;

Query Match 100.0%; Score 21; DB 2; Length 822;
 Best Local Similarity 100.0%; Pred. No. 1.9;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAGGACACAGTGGCTTTGAC 21
 |||||
 Db 676 CAGGACACAGTGGCTTTGAC 656

RESULT 41
 AED52633/C
 ID AED52633 standard; DNA; 822 BP.
 XX
 AC AED52633;
 XX
 DT 29-DEC-2005 (first entry)
 XX

DE Fusion protein D1/3-B6-H1s/HPV16, DNA.
 XX
 KM Fusion protein; vaccine; papilloma; cytosolic; papillomavirus infection;
 XX virucide; uterine cervix tumor; B6; ds; gene; D protein.
 XX
 OS Haemophilus influenzae; strain 772.
 OS Human papillomavirus type 16.
 OS Synthetic.
 OS Chimeric.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..822
 FT /*tag= a
 FT /product= "Fusion protein D1/3-B6-H1s/HPV16"
 XX
 IN9801903-14.
 XX
 PD 04-MAR-2005.
 XX
 PF 24-AUG-1998; 98IN-CH001903.
 XX
 PR 22-AUG-1997; 97BP-00179535.
 XX
 PA (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
 XX
 PI Tyrrell AMR;
 XX
 DR WPI; 2005-557648/57.
 DR P-PSDB; AED52634.
 XX
 PT Vaccine.
 XX
 PS Example 4; Fig 3; 96pp; English.
 XX
 CC The invention relates to human Papilloma virus (HPV) fusion proteins,
 CC linked to an immunological fusion partner that provides T helper epitopes
 CC to the HPV antigen. Vaccine formulations comprising the fusion proteins
 CC are useful in the treatment or prophylaxis of HPV induced lesions
 CC (papillomas, warts and cervical cancer). HPV 16 (and HPV 18) B6 and E7
 CC proteins (singly, as an B6-E7 fusion or mutated) were fused to either
 CC Haemophilus influenzae D protein (20-127), the C-terminus of
 CC Streptococcus pneumoniae LytA protein (cLYtA) or thiorodoxin. The present
 CC sequence encodes an HPV-H. influenzae D protein, fusion protein of the
 CC invention.
 CC
 SQ Sequence 822 BP; 277 A; 147 C; 168 G; 230 T; 0 U; 0 Other;

Query Match 100.0%; Score 21; DB 14; Length 822;
 Best Local Similarity 100.0%; Pred. No. 1.9;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAGGACACAGTGGCTTTGAC 21
 |||||
 Db 676 CAGGACACAGTGGCTTTGAC 656

RESULT 42
 AAX78795/C
 ID AAX78795 standard; DNA; 879 BP.
 XX
 AC AAX78795;
 XX
 DT 06-SEP-1999 (first entry)
 XX
 DE HPV fusion protein CLYTA-B6-H1s/HPV16 DNA.
 XX
 KM Fusion protein; B6 protein; E7 protein; B6/E7; immunomodulator; tumour;
 KM immunological fusion partner; Cpg oligonucleotide; immune response;
 KM HPV antigen; prevention; treatment; ss.
 XX
 OS Synthetic.
 OS Human papillomavirus.
 XX

```

PN WO9933868-A2.
XX
XX 08-JUL-1999.
XX
XX 18-DEC-1998; 98MO-EP008563.
XX
XX 24-DEC-1997; 97GB-00027262.
XX
XX (SMIK ) SMITHKLIN BEERHAM BIOLOGICALS.
XX
XX Dalemans WLJ, Gerard CMG;
XX
XX WPI, 1999-405485/34.
XX
XX P-PSDB; AAY25379.
XX
XX Composition comprising an E6, E7 or E6/E7 fusion protein from HPV to
XX induce immune response to HPV.
XX
XX Example VI, Page 52; 62pp; English.
XX
XX AAY78791-X78801 represent nucleic acid sequences which encode novel
XX constructs comprising an E6 or E7 protein or E6/E7 fusion protein from
XX HPV (represented in AAY25375-Y25386). These constructs are optionally
XX linked to an immunological fusion partner and an immunomodulatory CpG
XX oligonucleotide. The products of the invention can be used to induce an
XX immune response in a patient to an HPV antigen. They can also be used for
XX preventing or treating HPV induced tumours
XX
XX Sequence 879 BP; 289 A; 168 C; 209 G; 213 T; 0 U; 0 Other;
XX
XX Query Match 100.0%; Score 21; DB 2; Length 879;
XX Best Local Similarity 100.0%; Pred. No. 1.9;
XX Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 CAGGACACAGTGGCTTTTGAC 21
XX 733 CAGGACACAGTGGCTTTTGAC 713
XX
XX RESULT 43
XX AAY29784/C
XX ID AAY29784 standard; DNA; 879 BP.
XX
XX AAY29784;
XX
XX 17-OCT-2003 (revised)
XX DT 22-JUN-1999 (first entry)
XX
XX CLYTA-B6-His coding sequence.
XX
XX Chimeric; E6; E7; fusion protein; CLYTA; vaccine; immunotherapy; tumour;
XX lesion; benign; malignant; virus; infection; ss.
XX
XX Human papillomavirus.
XX OS Streptococcus pneumoniae.
XX Chimeric.
XX
XX WO9910375-A2.
XX
XX 04-MAR-1999.
XX
XX 17-AUG-1998; 98MO-BP005285.
XX
XX 22-AUG-1997; 97GB-00017953.
XX
XX (SMIK ) SMITHKLIN BEERHAM BIOLOGICALS.
XX
XX Bruck C, Cabazon Silva T, Delisse ABF, Gerard CMG;
XX Lombardo-Bencheikh A;
XX
XX WPI, 1999-190587/16.
XX
XX P-PSDB; AAY02635.
XX

```

```

PT Human Papilloma Virus (HPV) fusion proteins - useful in vaccines for
XX treatment or prophylaxis of HPV induced lesions.
XX
XX Disclosure; Fig 10; 95pp; English.
XX
XX This sequence represents the coding region for a chimeric E6 or E7
XX protein or E6/E7 fusion protein from Human papillomavirus (HPV) linked to
XX an immunological fusion partner. In this case, a fragment of the
XX Streptococcus pneumoniae CLYTA protein. The sequence also contains a
XX histidine tag at the C-terminus of the encoded protein. The protein can
XX be used in a vaccine, for immuno-therapeutically treating HPV induced
XX tumour lesions (benign or malignant) and preventing HPV viral infection.
XX (Updated on 17-OCT-2003 to standardise OS field)
XX
XX Sequence 879 BP; 289 A; 168 C; 209 G; 213 T; 0 U; 0 Other;
XX
XX Query Match 100.0%; Score 21; DB 2; Length 879;
XX Best Local Similarity 100.0%; Pred. No. 1.9;
XX Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 CAGGACACAGTGGCTTTTGAC 21
XX 733 CAGGACACAGTGGCTTTTGAC 713
XX
XX RESULT 44
XX AED52641/C
XX ID AED52641 standard; DNA; 879 BP.
XX
XX AED52641;
XX
XX 29-DEC-2005 (first entry)
XX DT
XX Fusion protein clyta-B6-His/HPV16, DNA.
XX
XX Fusion protein; vaccine; papilloma; cytostatic; papillomavirus infection;
XX virucide; uterine cervix tumor; E6; ds; gene; lyta.
XX
XX Streptococcus pneumoniae.
XX OS Human papillomavirus type 16.
XX Synthetic.
XX Chimeric.
XX
XX Key Location/Qualifiers
XX FH 1..879
XX CDS /*tag= a
XX FT /product= "Fusion protein clyta-B6-His/HPV16"
XX
XX IN9801903-14.
XX
XX 04-MAR-2005.
XX
XX 24-AUG-1998; 98IN-CH001903.
XX
XX 22-AUG-1997; 97BP-00179535.
XX
XX (SMIK ) SMITHKLIN BEERHAM BIOLOGICALS.
XX
XX Tyrtell AMR;
XX
XX WPI, 2005-557648/57.
XX
XX P-PSDB; AED52642.
XX
XX Vaccine.
XX
XX Example 10; Fig 10; 96pp; English.
XX
XX The invention relates to human Papilloma virus (HPV) fusion proteins,
XX CC linked to an immunological fusion partner that provides T helper epitopes
XX CC to the HPV antigen. Vaccine formulations comprising the fusion proteins
XX CC are useful in the treatment or prophylaxis of HPV induced lesions
XX CC (papillomas, warts and cervical cancer). HPV 16 (and HPV 18) E6 and E7
XX CC proteins (singly, as an E6-E7 fusion or mutated) were fused to either

```


CC Haemophilus influenzae D protein (20-127), the C-terminus of
CC Streptococcus pneumoniae LytA protein (cLytA) or thiorodoxin. The present
CC sequence encodes an HPV-LytA, fusion protein of the invention.

XX SQ Sequence 879 BP; 289 A; 168 C; 209 G; 213 T; 0 U; 0 Other;

Query Match 100.0%; Score 21; DB 14; Length 879;

Best Local Similarity 100.0%; Pred. No. 1.9;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAGGACACAGTGGCTTTGAC 21

Db 733 CAGGACACAGTGGCTTTGAC 713

RESULT 45
AAA09477/c

ID AAA09477 standard; DNA; 921 BP.

XX AAA09477;

AC 29-AUG-2000 (first entry)

XX Human papillomavirus type 16 E6 and E7 open reading frames.

XX hINP; involucrin; regulatory region; tissue-specific; promoter; oncogene;
XX suprabasal; stratified; squamous; epithelial tissue; animal model;
XX neoplasia; papillomavirus 16; E6; E7; blistering; proliferative;
XX ichthyosis; cancer; ss.

OS Human papillomavirus type 16.

XX Key Location/Qualifiers

FT CDS 170..625

FT /*cag= a

FT /label= B6

FT mat_peptide 628..921

FT /*cag= b

FT /label= E7

XX MO200026343-A2.

XX 11-MAY-2000.

XX 29-OCT-1999; 99WO-US025516.

XX 30-OCT-1998; 98US-0106495P.

XX 29-OCT-1999; 99US-00430201.

XX (UYCA-) UNIV CASE WESTERN RESERVE.

XX Eckert RL, Crish JF;

XX WPI; 2000-365596/31.

XX Transgenic animals that express oncogenes in epithelial tissue, useful as
XX disease models for neoplastic conditions such as psoriasis and melanomas.

XX Disclosure; Fig 13; 86pp; English.

XX The invention concerns transgenic non-human animals that express
XX oncogenes of interest in their epithelial cells, and materials and
XX methods for their production. Expression of the oncogenes is controlled
XX by a sequence from the human involucrin upstream regulatory region (see
XX AAA09476) which targets expression to desired tissues and cells (i.e.
XX suprabasal stratified squamous epithelial tissue). The transgenic animals
XX are useful as animal models for studying neoplastic diseases that affect
XX epithelial tissues, such as those caused by the papillomavirus 16
XX oncogene E6 or E7 nucleic acid sequences and including blistering
XX diseases (e.g. epidermolytic hyperkeratosis, Dowling-Meara disease),
XX proliferative diseases (e.g. psoriasis, epidermal dysplasia and Bulosa
XX simplex), ichthyosis disease (e.g. Ichthyosis bullosa Siemens and
XX recessive X-linked ichthyosis) and cancers of the skin, breast, lung,

CC colon, cervix and liver. The animal models may also be used to screen
CC candidate agents for the ability to prevent or treat these diseases
CC (claimed)

XX SQ Sequence 921 BP; 308 A; 167 C; 197 G; 249 T; 0 U; 0 Other;

Query Match 100.0%; Score 21; DB 3; Length 921;

Best Local Similarity 100.0%; Pred. No. 1.9;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAGGACACAGTGGCTTTGAC 21

Db 506 CAGGACACAGTGGCTTTGAC 486

RESULT 46
AAD35101/c

ID AAD35101 standard; DNA; 939 BP.

XX AAD35101;

AC 25-JUL-2002 (first entry)

XX Human papillomavirus enh6,7 construct DNA.

XX Alphavirus vector system; human papilloma virus; HPV; cervical cancer;
XX therapy; vaccine; virucide; cytostatic; ds.
XX Human papillomavirus.

OS EPI195438-A1.

XX 10-APR-2002.

XX 06-OCT-2000; 2000EP-00203472.

XX 06-OCT-2000; 2000EP-00203472.

XX (UYGR-) RIJSDUNIV GRONINGEN.

XX Regte DG, Wilschut JC, Holtrop M, Daemen CAHH;

XX WPI; 2002-354156/39.

XX New alphavirus system, useful for genetic immunization against cervical
XX cancer, comprises papilloma virus nucleic acid.

XX Example 2; Fig 19; 45pp; English.

XX The present invention relates to an alphavirus vector system comprising
XX nucleic acid derived from a human papilloma virus (HPV). The invention or
XX cells containing it, are used in treatment and prevention of cervical
XX cancer, particularly as a vaccine. By selecting the nucleic acid that
XX encode E6/E7 proteins without ability to bind to pRb and p53, the risk
XX that cells infected with the alphavirus vector system may become
XX oncogenic is avoided (contrast use of other viral vectors). The present
XX sequence is Human papillomavirus enh6,7 construct DNA

XX SQ Sequence 939 BP; 284 A; 197 C; 217 G; 241 T; 0 U; 0 Other;

Query Match 100.0%; Score 21; DB 6; Length 939;

Best Local Similarity 100.0%; Pred. No. 1.9;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAGGACACAGTGGCTTTGAC 21

Db 522 CAGGACACAGTGGCTTTGAC 502

RESULT 47

AAQ08627/c

ID AAQ08627 standard; DNA; 1000 BP.

XX

```

AC  AAQ08627;
XX
XX  24-OCT-2003 (revised)
DT  21-APR-1994 (first entry)
XX
XX  HPV-16 fragment.
DE
XX  HPV-18; HPV-16; amplification; primer; polymerase chain reaction; PCR;
KM  ss.
XX
XX  Human papillomavirus type 16.
OS
XX
XX  Key
FH  Location/Qualifiers
FT  misc_binding
   /tag= a
   /note= "primer (AAQ08628) binding site"
FT  misc_binding
   /tag= b
   /note= "primer (AAQ08629) binding site"
FT  misc_binding
   /tag= c
   /note= "primer (AAQ08630) binding site"
XX
XX  DE3838269-A.
XX
XX  17-MAY-1990.
XX
XX  11-NOV-1988; 88DE-03838269.
XX
XX  11-NOV-1988; 88DE-03838269.
XX
XX  (BEHW ) BEHRINGER AG.
XX
XX  Cerutti P, Whitcomb J, Zijlstra J, Devilliers EM;
PI  WPI, 1990-156905/21.
XX
XX  Detection of human papilloma virus - by DNA amplification and analysis.
XX
XX  Example 4b; Page 4; 11pp; German.
XX
XX  Example 4b describes the results of the amplification of HPV-16 DNA by
CC  PCR using primers. (Updated on 24-OCT-2003 to standardise OS field)
XX
XX  Sequence 1000 BP; 340 A; 171 C; 230 G; 259 T; 0 U; 0 Other;
SQ
XX
XX  Query Match 100.0%; Score 21; DB 2; Length 1000;
XX  Best Local Similarity 100.0%; Pred. No. 1.9;
XX  Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY
   1 CAGGACACAGTGGCTTTTGAC 21
   |||||
DB  440 CAGGACACAGTGGCTTTTGAC 420

RESULT 48
AA91784/c
ID  AA91784 standard; DNA; 1005 BP.
XX
XX  AA91784;
AC
XX
XX  27-AUG-2003 (revised)
DT  25-MAR-2003 (revised)
DT  16-MAR-1990 (first entry)
XX
XX  DNA probe complementary to human papilloma virus type 16.
XX  Cervical cancer.
XX
XX  Human papillomavirus type 16.
XX  MO8909940-A.
XX

```

```

PD  19-OCT-1989.
XX
XX  04-APR-1989; 89MO-US001318.
XX
XX  04-APR-1988; 88US-00177404.
XX
XX  31-MAR-1989; 89US-00330381.
XX
XX  (ONCO-) ONCOR INC.
XX
XX  George AL, Groff DE;
PI  WPI, 1989-324314/44.
XX
XX  Rapid detection of specific human papilloma virus genotypes - by
PT  hybridisation of DNA digest with new labelled nucleic acid probes.
XX
XX  Claim 40; Page 46; 81pp; English.
XX
XX  Obt. by cutting HPV16 with BamHI and PvuII. The patent describes probes
CC  (DNA or RNA) and their complements capable of detecting one or a
CC  combination of HPV types 6, 11, 16, 18, 31, 33 and 35. (Updated on 25-MAR
CC  -2003 to correct PR field.) (Updated on 27-AUG-2003 to correct OS field.)
XX
XX  Sequence 1005 BP; 325 A; 182 C; 190 G; 308 T; 0 U; 0 Other;
SQ
XX
XX  Query Match 100.0%; Score 21; DB 1; Length 1005;
XX  Best Local Similarity 100.0%; Pred. No. 1.9;
XX  Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY
   1 CAGGACACAGTGGCTTTTGAC 21
   |||||
DB  892 CAGGACACAGTGGCTTTTGAC 872

RESULT 49
AA878793/c
ID  AA878793 standard; DNA; 1116 BP.
XX
XX  AA878793;
AC
XX
XX  06-SRP-1999 (first entry)
DT
XX
XX  HPV fusion protein DI/3-E6E7-His/HPV16 DNA.
DE
XX
XX  Fusion protein; E6 protein; E7 protein; E6/E7; Immunomodulator; tumour;
KM  immunological fusion partner; CPG oligonucleotide; immune response;
KM  HIV antigen; prevention; treatment; ss.
XX
XX  Synthetic.
OS  Human papillomavirus.
XX
XX  MO9933868-A2.
XX
XX  08-JUL-1999.
PD
XX
XX  18-DEC-1998; 98WO-EP008563.
XX
XX  24-DEC-1997; 97GB-00027262.
XX
XX  (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.
XX
XX  Dalemans WLJ, Gerard CMG;
PI  WPI, 1999-405485/34.
XX
XX  P-PsDB; AAY25377.
DR
XX
XX  Composition comprising an E6, E7 or E6/E7 fusion protein from HPV to
PT  induce immune response to HPV.
XX
XX  Example III; Page 49; 62pp; English.
XX
XX  AA878791-X78801 represent nucleic acid sequences which encode novel
CC  constructs comprising an E6 or E7 protein or E6/E7 fusion protein from

```

CC HPV (represented in AAY25375-Y25386). These constructs are optionally
 CC linked to an immunological fusion partner and an immunomodulatory Cpg
 CC oligonucleotide. The products of the invention can be used to induce an
 CC immune response in a patient to an HPV antigen. They can also be used for
 CC preventing or treating HPV induced tumours
 XX

Db 676 CAGGACACAGTGGCTTTGAC 656
 Search completed: May 24, 2006, 05:56:59
 Job time : 256.028 secs

SQ Sequence 1116 BP; 368 A; 208 C; 235 G; 305 T; 0 U; 0 Other;
 Query Match 100.0%; Score 21; DB 2; Length 1116;
 Best Local Similarity 100.0%; Pred. No. 2;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAGGACACAGTGGCTTTGAC 21
 |||||
 Db 676 CAGGACACAGTGGCTTTGAC 656

RESULT 50
 AAX29782/c
 ID AAX29782 standard; DNA; 1116 BP.
 XX

AC AAX29782;
 XX
 DT 17-OCT-2003 (revised)
 DT 22-JUN-1999 (first entry)
 XX

DE Prot.D1/3-E6-E7-His/HPV16 coding sequence.

XX Chimeric; E6; E7; fusion protein; protein D; vaccine; immunotherapy;
 KM tumour; lesion; benign; malignant; virus; infection; ss.
 XX

OS Human Papillomavirus.
 OS Haemophilus influenzae.
 OS Chimeric.

XX WO910375-A2.

XX 04-MAR-1999.

XX 17-AUG-1998; 98WO-EP005285.

XX 22-AUG-1997; 97GB-00017953.

XX (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.

XX Bruck C, Cabezon Silva T, Delisse AEF, Gerard CMG;
 PI Lombardo-Bencheikh A;
 XX

XX WPI; 1999-190587/16.
 DR P-PSDB; AAY02633.
 XX

PT Human Papilloma Virus (HPV) fusion proteins - useful in vaccines for
 PT treatment or prophylaxis of HPV induced lesions.
 XX

PS Disclosure; Fig 6; 95pp; English.

XX This sequence represents the coding region for a chimeric E6 or E7
 CC protein or E6/E7 fusion protein from Human papillomavirus (HPV) linked to
 CC an immunological fusion partner, in this case, a fragment of the
 CC Haemophilus influenzae B protein D. The sequence also contains a
 CC histidine tag at the C-terminus of the encoded protein. The protein can
 CC be used in a vaccine, for immuno-therapeutically treating HPV induced
 CC tumour lesions (benign or malignant) and preventing HPV viral infection.
 CC (Updated on 17-OCT-2003 to standardise OS field)
 CC

XX Sequence 1116 BP; 368 A; 208 C; 235 G; 305 T; 0 U; 0 Other;

Query Match 100.0%; Score 21; DB 2; Length 1116;
 Best Local Similarity 100.0%; Pred. No. 2;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAGGACACAGTGGCTTTGAC 21
 |||||

THIS PAGE IS BLANK

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 24, 2006, 05:47:01 ; Search time 1449.58 Seconds
(without alignments)
810.102 Million cell updates/sec

Title: US-10-601-913-85

Perfect score: 21

Sequence: 1 CAGGACACAGTGGCTTTTAC 21

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 48236798 seqs, 27959665780 residues

Total number of hits satisfying chosen parameters: 96473596

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 1000 summaries

Database :

EST:
1: gb_est1:*
2: gb_est3:*
3: gb_est4:*
4: gb_est5:*
5: gb_est6:*
6: gb_est7:*
7: gb_est8:*
8: gb_est9:*
9: gb_est10:*
10: gb_est11:*
11: gb_est12:*
12: gb_est13:*
13: gb_est14:*
14: gb_est15:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	21	100.0	244	3	BO106475
2	18.4	87.6	229	7	BB264006
3	18.4	87.6	291	7	BB264006
4	18.4	87.6	291	7	BB264006
5	18.4	87.6	534	8	CR538057
6	18.4	87.6	546	9	DA514546
7	18.4	87.6	549	9	DA722798
8	18.4	87.6	583	9	DA906377
9	18.4	87.6	644	7	BB647218
10	18.4	87.6	654	13	C2303586
11	18.4	87.6	660	5	CJ337277
12	18.4	87.6	720	5	CD353955
13	18.4	87.6	2148	14	AY406159
14	18.4	87.6	2400	14	AY406158
15	18.4	87.6	2400	14	AY406160
16	18.4	87.6	2711	6	AK139334
17	18.4	87.6	3161	6	AK044046
18	18.4	87.6	3592	6	AK083092
19	18.4	87.6	3793	6	AK086711

20	17.8	84.8	337	2	BG022939	BG022939 daa81d07.
21	17.8	84.8	357	2	BF771220	BF771220 IL5-IT002
22	17.8	84.8	377	2	CR555978	CR555978 DKFPP4691
23	17.8	84.8	501	4	BY467982	BY467982 BY467982
24	17.8	84.8	507	8	CR429603	CR429603 CR429603
25	17.8	84.8	594	1	AL785427	AL785427 AL785427
26	17.8	84.8	595	9	CK421278	CK421278 JGI_XZG15
27	17.8	84.8	626	14	DR5G15T	AL733484 Danilo rer
28	17.8	84.8	643	11	AQ65517	AQ625517 CIRIB1-RL
29	17.8	84.8	649	1	AL642415	AL642415 AL642415
30	17.8	84.8	683	14	CT346911	CT346911 Sub scrofi
31	17.8	84.8	710	14	AG144458	AG144458 Pan trogl
32	17.8	84.8	712	11	A2560569	A2560569 RPCI-23-2
33	17.8	84.8	788	9	CK380870	CK380870 JGI_XZT52
34	17.8	84.8	793	9	CK443805	CK443805 JGI_XZG94
35	17.8	84.8	794	9	CK444795	CK444795 JGI_XZG10
36	17.8	84.8	796	9	CK404457	CK404457 JGI_XZT35
37	17.8	84.8	799	9	CK439900	CK439900 JGI_XZG80
38	17.8	84.8	829	9	CK407260	CK407260 JGI_XZT62
39	17.8	84.8	850	9	CK396192	CK396192 JGI_XZT56
40	17.8	84.8	874	12	CC295265	CC295265 CH261-52A
41	17.8	84.8	883	9	CK418837	CK418837 JGI_XZG65
42	17.8	84.8	895	10	DR897380	DR897380 JGI_XZT42
43	17.8	84.8	895	10	DT398630	DT398630 JGI_CAB12
44	17.8	84.8	896	8	CK333579	CK333579 JGI_XZT69
45	17.8	84.8	906	5	CP222676	CP222676 AGENCOURT
46	17.8	84.8	971	12	CC304636	CC304636 CH261-89N
47	17.8	84.8	2617	6	AK034662	AK034662 Mus muscu
48	17.8	84.8	5736	6	AK157580	AK157580 Mus muscu
49	17.8	82.9	390	12	CE693411	CE693411 tigr-g88-
50	17.4	82.9	430	11	AQ435575	AQ435575 HS_5147 A
51	17.4	82.9	472	4	CB730452	CB730452 AMGNNUC-N
52	17.4	82.9	503	14	CR266963	CR266963 Forward s
53	17.4	82.9	545	1	AJ940900	AJ940900 AJ940900
54	17.4	82.9	602	8	CV429532	CV429532 52239.1 A
55	17.4	82.9	630	2	BC083368	BC083368 H3087B03-
56	17.4	82.9	739	13	CZ080667	CZ080667 OM_Ba007
57	17.4	82.9	753	1	AJ943303	AJ943303 AJ943303
58	17.4	82.9	792	8	CV450533	CV450533 EST904442
59	17.4	82.9	981	12	CC192876	CC192876 CH261-35F
60	17.4	82.9	1197	12	CC287367	CC287367 CH261-34D
61	17.4	82.9	246	7	BB412072	BB412072 BB412072
62	17.4	81.0	408	3	BP715091	BP715091 BP715091
63	17.4	81.0	469	10	DT926438	DT926438 BMSB10674
64	17.4	81.0	480	10	DT920674	DT920674 BMSB11237
65	17.4	81.0	492	7	BF049136	BF049136 dc04g05.x
66	17.4	81.0	494	2	BS579815	BS579815 dab77e04.
67	17.4	81.0	497	2	BJ079209	BJ079209 BJ079209
68	17.4	81.0	515	11	AQ318368	AQ318368 RPCI11-10
69	17.4	81.0	542	10	DM550648	DM550648 EST_g881-
70	17.4	81.0	603	3	BP724122	BP724122 BP724122
71	17.4	81.0	635	3	BU419446	BU419446 603961749
72	17.4	81.0	637	3	BP724451	BP724451 BP724451
73	17.4	81.0	643	3	BP718952	BP718952 BP718952
74	17.4	81.0	652	3	BU448384	BU448384 603213573
75	17.4	81.0	677	2	BT094427	BT094427 BT094427
76	17.4	81.0	678	10	DV868100	DV868100 CRE7832 C
77	17.4	81.0	680	2	BU075913	BU075913 BU075913
78	17.4	81.0	690	2	BP722067	BP722067 BP722067
79	17.4	81.0	700	4	CB600923	CB600923 AGENCOURT
80	17.4	81.0	716	2	BU632398	BU632398 BU632398
81	17.4	81.0	750	2	BU644127	BU644127 BU644127
82	17.4	81.0	757	3	BP701906	BP701906 BP701906
83	17.4	81.0	769	10	DR721880	DR721880 AGENCOURT
84	17.4	81.0	797	3	BP683846	BP683846 BP683846
85	17.4	81.0	798	3	BP721999	BP721999 BP721999
86	17.4	81.0	821	3	BP719880	BP719880 BP719880
87	17.4	81.0	857	4	CB199693	CB199693 AGENCOURT
88	17.4	81.0	861	3	BP733643	BP733643 BP733643
89	17.4	81.0	863	3	BP734862	BP734862 BP734862
90	17.4	81.0	871	4	CB564030	CB564030 AGENCOURT
91	17.4	81.0	896	4	CB561058	CB561058 AGENCOURT
92	17.4	81.0	926	4	CB560615	CB560615 AGENCOURT

969	15.8	75.2	384	11	B366210
970	15.8	75.2	386	11	B90843
971	15.8	75.2	388	11	AM900115
972	15.8	75.2	389	1	A1100673
973	15.8	75.2	391	3	BUT73386
974	15.8	75.2	391	7	AM665451
975	15.8	75.2	397	12	B2938673
976	15.8	75.2	398	2	BI473695
977	15.8	75.2	398	11	B69790
978	15.8	75.2	399	1	A1035861
979	15.8	75.2	402	1	AA602631
980	15.8	75.2	403	11	B34265
981	15.8	75.2	404	4	BY279842
982	15.8	75.2	404	8	CO208022
983	15.8	75.2	406	2	BM228285
984	15.8	75.2	409	1	A1458745
985	15.8	75.2	409	2	BM225622
986	15.8	75.2	409	11	B91445
987	15.8	75.2	411	2	BM193784
988	15.8	75.2	412	1	AA045612
989	15.8	75.2	412	1	AV597884
990	15.8	75.2	413	3	BO670910
991	15.8	75.2	415	3	BG799187
992	15.8	75.2	415	3	BO001927
993	15.8	75.2	416	3	BP786395
994	15.8	75.2	416	11	AQ185395
995	15.8	75.2	417	7	AM338111
996	15.8	75.2	418	11	AQ0825187
997	15.8	75.2	419	7	AM450167
998	15.8	75.2	422	1	AA378780
999	15.8	75.2	424	1	AA895740
C1000	15.8	75.2	424	1	AA402693
					B366210 CH230-170
					B90843 C1T-HSP-217
					AM900115 CM0-NN100
					A1100673 35048 Lam
					BUT73386 UT-E-C11-
					AM665451 b104403.x
					B2938673 CH240-64C
					BI473695 fP41A11.Y
					B69790 C1T-HSP-205
					A1035861 uD46h11.x
					AA602631 nO91605.s
					B34265 HS-1024-A2-
					BY279842 B279842
					CO208022 WS0094.B2
					BM228285 K0258E03-
					A1458745 C197f11.x
					BM225622 K0217G07-
					B91445 C1T-HSP-217
					BM193784 TCAAP1D82
					AA045612 zK83a10.s
					AV597884 AV597884
					BO670910 AGENCOURT
					BG799187 fP31d05.Y
					BO001927 UT-H-DH1-
					BP786395 BP786395
					AQ185395 HS J138-B
					AM338111 xw66h03.x
					AQ0825187 HS 5254.A
					AM450167 UT-H-B13-
					AA378780 B8T91515
					AA895740 y114d06.L
					AA402693 z147d05.s

ALIGNMENTS

BH366210 CT CH230-12.7
 B90843 CIT-HSP-12.7
 AM900115 CM0-NN100
 AI100673 35048 Lam
 BU733286 UT-E-CIT-1
 AM665451 h104a03.3
 B23384673 CH240.66g
 B1473695 IP41a1-1
 B69790 CT-HSP-1205
 A1035861 na46hm1.r
 A1003661 nc91ne05.s
 B34265 H5-1024-A2-5
 B2789842 B2798942
 CO200822 WS0094.B2
 BM228285 CO208803
 AI158745 t397a11.x
 BM228822 CT-H217607-
 B91445 CIT-HSP-12.7
 BM195674 TC0A01PD82
 AA0495612 z183a10.8
 A15927884 A15927884
 B6770910 AGNCOYRT
 BG961048 f531d05.Y
 Q000101927 UT-H-DH1-
 BP786395 BP786395
 AQ1883334 HS 3138.B
 A1381111 xw66n03.x
 AA0502167 UT-H-5254.A
 A4085167 UT-H-815
 AA378780 B59191515
 AA085740 vyi4a05.b
 AA402693 z147a05.8

FEATURES	source
LOCUS	244 bp mRNA linear EST 16-APR-2000
DEFINITION	fc2102.e Rose Petals (Fragrant Cloud) Lambda Zap Express Library
ACCESSION	BQ106475
VERSION	BQ106475
KEYWORDS	BQ106475.1 GI:20156137
SOURCE	EST.
ORGANISM	Rosa hybrid cultivar
REFERENCE	Rosa hybrid cultivar
AUTHORS	Eukaryotes, Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophytes; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosid 1; Rosales; Rosaceae; Rosoideae; Rosa.
	1 (bases 1 to 244)
	Guterman,I., Shalit,M., Menda,N., Pleskun,D., Dafny-Yelin,M., Shalev,G., Bar,E., Davidov,O., Ovadis,M., Emanuel,M., Wang,J., Adam,Z., Pichersky,R., Lewinson,E., Zamir,D., Vainstein,A. and Weiss,D.
TITLE	Rose Scent: Genomics Approach to Discovering Novel Floral
JOURNAL	Fragrance-Related Genes
COMMENT	Plant Cell 14 (10), 2325-2338 (2002)
	Contact: Naama Menda
	Petal Genomics
	Faculty of Agricultural, Food and Environmental Quality Sciences,
	The Hebrew University of Jerusalem
	P.O. Box 12, Rehovot, 76100, Israel
	Tel: 972 8 9489 683
	Fax: 972 8 9468 265
	Email: shaham@rl.huji.ac.il
	Seq primer: 73 forward.
	Location/Qualifiers
	1..244

```
FEATURES
  source
    location/Qualifiers
      1. 244
        /organism="Rosa hybrid cultivar"
        /mol_type="mRNA"
        /strain="Fragrant Cloud"
        /db_xref="taxon:128735"
```

```

/clone="fc2102.e"
/tissue_type="Petals"
/dev_stage="Young open flower at stage four"
/clone_id="Rose Petals (Fragment Cloud) Lambda Zap Express Library"
/note="Vector: pBKCMV; Site_1: EcoRI; Site_2: XhoI"

ORIGIN

Query Match      100.0%; Score 21; DB 3; Length 244;
Best Local Similarity 100.0%; Prid. No. 35;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      1 CAGGACACAGTGGCTTTTGAC 21
         |||||
Db       121 CAGGACACAGTGGCTTTTGAC 101

```

RESULT 2				
B833633	B833633	229 bp	mRNA	linear EST_11-JUL-2000
LOCUS	B833633			
DEFINITION	B833633 RIKEN full-length enriched, 10 days neonate medulla oblongata Mus musculus cDNA clone B830014D01.3, mRNA sequence.			
ACCESSION	B833633			
VERSION	B833633.1	GI:9042396		
KEYWORDS	EST.			
SOURCE	Mus musculus (house mouse)			
ORGANISM	Mus musculus			
	Eumetazoa; Metazoa; Chordata; Craniata; Vertebrate; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciuromorphi; Muroidae; Muridae; Murinae; Mus.			
REFERENCE	1 (bases 1 to 229)			
AUTHORS	Kondo,H., Aizawa,K., Akahira,S., Akiyama,T., Arakawa,T.,			

TITLE	JOURNAL	COMMENT
Unpublished (2000)		
Contact: Yoshihide Hayashizaki		
Hayashizaki Y.		
RIKEN Mouse ESTs (Kono, H., et al.)		
Kiyosawa, H., Kojima, Y., Kondo, S., Koye, S., Kurinaka, C., Kutsakebe, M., Matsuyama, T., Miki, R., Mizuno, Y., Nakamura, M., Oda, H., Okazaki, Y., Ono, T., Owa, C., Saito, H., Sakai, C., Sato, K., Shibata, K., Shibata, Y., Shigemoto, Y., Shingawa, A., Shiraki, T., Sogabe, Y., Sugahara, Y., Suzuki, H., Suzuki, H., Tagawa, A., Takahashi, F., Tomimaga, N., Toyota, T., Tsunoda, Y., Watanishi, A., Watanabe, S., Yamamura, T., Yamana, T., Yano, R., Yasunishi, A., Yokota, T., Yoshida, K., Yoshiki, A., Yoshino, M., Muramatsu, M. and Hayashizaki Y.		

The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@isc.riken.jp, URL: <http://genome.gsc.riken.jp/>
Carninci, P., Nishiyama, Y., Westover, A., Itoh, M., Nagaoaka, S.,
Saeki, N., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.
Thermotranscription and thermocatalysis of thermolabile enzymes by
transposon and its application for the synthesis of full length
cDNA. *Proc. Natl. Acad. Sci. U.S.A.* 95 (2): 520-524 (1998)
Itoh, M., Kikunishi, T., Akiyama, J., Shibata, K., Izawa, M., Kawai, J.,
Tomaru, Y., Carninci, P., Shibata, Y., Ozawa, Y., Muramatsu, M.,
Okazaki, Y., and Hayashizaki, Y.
Automated filtration-based high-throughput plasmid preparation
system. *Genome Res.* 9 (5): 463-470 (1999)
Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning. *Methods Enzymol.* 303,
19-44 (1999)
Please visit our web site (<http://genome.irc.riken.go.jp>) for
further details.

location/Qualifiers

1. 229

organism="Mus musculus"
mol_type="mRNA"

/db xref="taxon:10090"
 /clone="B830014D01"
 /tissue_type="medulla oblongata"
 /dev_stage="10 days neonate"
 /lab_host="DH10B"
 /clone_1ib="RIKEN full-length enriched, 10 days neonate medulla oblongata"
 /note="Site 1: SalI; Site 2: BamHI. cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5']
 GAGAGAGAGAGATCCAGAGCTCTTTTCTTTTCTTTTCTTTT 3', cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 20.0 and subtraction to Rot = 459.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGATTCGAGTTATTAATTATATCCCCCCCC 3']. cDNA was cleaved with XhoI and BamHI. Vector: a modified pluescript KS(+) after bulk excision from Lambda Flc I."

ORIGIN

Query Match 87.6%; Score 18.4; DB 7; Length 229;
 Best Local Similarity 95.0%; Pred. No. 6.6e+02;
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AGGACACAGTGGCTTTTGAC 21
 Db 93 AGGACACAGTGGCTTTTGAC 112

RESULT 3
 BB264006 291 bp mRNA linear EST 07-JUL-2000
 LOCUS BB264006 RIKEN full-length enriched, 10 days neonate cortex Mus
 DEFINITION Musculus cDNA clone AB30013L24 3' similar to X95600 M.musculus mRNA for cadherin-8, mRNA sequence.

ACCESSION BB264006.1 GI:8960462
 VERSION BB264006
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus

REFERENCE
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridea; Muridae; Murinae; Mus.

1 (bases 1 to 291)
 Komno, H., Aizawa, K., Akahira, S., Akiyama, J., Arakawa, T., Carninci, P., Endo, T., Fukuda, S., Fukunishi, Y., Hara, A., Hayatsu, N., Hirozane, T., Hori, F., Ishii, Y., Ishikawa, J., Ishikawa, T., Itoh, M., Izawa, M., Kadota, K., Kagawa, I., Kai, C., Kawai, J., Kikuchi, N., Kiyosawa, H., Kojima, Y., Kondo, S., Koya, S., Kurihara, C., Kuwabake, M., Matsuyama, T., Miki, R., Mizuno, Y., Nakamura, M., Oda, H., Okazaki, Y., Ono, T., Owa, C., Saito, H., Sakai, C., Sato, K., Shibata, K., Shibata, Y., Shigemoto, Y., Shinagawa, A., Shiraki, T., Sogabe, Y., Sugahara, Y., Suzuki, H., Suzuki, H., Tagawa, A., Takehashi, F., Tomioka, N., Toyota, T., Tsunoda, Y., Warahiki, A., Watanabe, S., Yamamura, T., Yamashita, I., Yano, R., Yasunishi, A., Yokota, T., Yoshida, K., Yoshiki, A., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.

RIKEN Mouse ESTs (Komno, H., et al.)
 Unpublished (2000)

TITLE
 JOURNAL
 COMMENT
 Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Suenho-cho, Tsunumi-ku, Yokohama, Kanagawa 230-0045, Japan
 Tel: 81-45-503-9222
 Fax: 81-45-503-9216
 Email: genome-res@gscc.riken.jp, URL: http://genome.gsc.riken.jp/

FEATURES

source

Carninci, P., Nishiyama, Y., Westover, A., Itoh, M., Nagaoka, S., Sasaki, N., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 Thermoactivation and thermocycling of thermolabile enzymes by trehalose and its application for the synthesis of full length cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)
 Itoh, M., Katsunari, T., Akiyama, J., Shibata, K., Izawa, M., Kawai, J., Tomaru, Y., Carninci, P., Shibata, Y., Ozawa, Y., Muramatsu, M., Okazaki, Y. and Hayashizaki, Y.
 Automated filtration-based high-throughput plasmid preparation system. Genome Res. 9 (5), 463-470 (1999)
 Carninci, P. and Hayashizaki, Y.
 High-efficiency full-length cDNA cloning. Methods Enzymol. 303, 19-44 (1999)
 Please visit our web site (http://genome.etc.riken.go.jp) for further details.

Location/Qualifiers

1..291

/organism="Mus musculus"

/mol_type="mRNA"

/db_xref="taxon:10090"

/clone="AB30013L24"

/tissue_type="cortex"

/dev_stage="10 days neonate"

/lab_host="DH10B"

/clone_1ib="RIKEN full-length enriched, 10 days neonate cortex"

/note="Site 1: SalI; Site 2: BamHI. cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5']
 GAGAGAGAGAGATCCAGAGCTCTTTTCTTTTCTTTTCTTTT 3', cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 20.0 and subtraction to Rot = 459.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGATTCGAGTTATTAATTATATCCCCCCCC 3']. cDNA was cleaved with XhoI and BamHI. Vector: a modified pluescript KS(+) after bulk excision from Lambda Flc I."

ORIGIN

Query Match 87.6%; Score 18.4; DB 7; Length 291;
 Best Local Similarity 95.0%; Pred. No. 6.8e+02;
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AGGACACAGTGGCTTTTGAC 21
 Db 150 AGGACACAGTGGCTTTTGAC 169

RESULT 4
 BB347346 291 bp mRNA linear EST 12-JUL-2000
 LOCUS BB347346 RIKEN full-length enriched, 10 days neonate cerebellum Mus
 DEFINITION Musculus cDNA clone B930065F09 3' similar to X95600 M.musculus mRNA for cadherin-8, mRNA sequence.

ACCESSION BB347346.1 GI:9059174
 VERSION BB347346
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus

REFERENCE
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridea; Muridae; Murinae; Mus.

1 (bases 1 to 291)
 Komno, H., Aizawa, K., Akahira, S., Akiyama, J., Arakawa, T., Carninci, P., Endo, T., Fukuda, S., Fukunishi, Y., Hara, A., Hayatsu, N., Hirozane, T., Hori, F., Ishii, Y., Ishikawa, J., Ishikawa, T., Itoh, M., Izawa, M., Kadota, K., Kagawa, I., Kai, C., Kawai, J., Kikuchi, N.,

2 AGGACACAGTGGCTTTTGAC 21

Yamashita, R., Yamamoto, J., Sekine, M., Tsuritani, K.,
Wakaguri, H.,
Ishii, S., Sugiyama, T., Saito, K., Isono, Y., Irle, R.,
Kushida, N.,

Yoneyama, T., Otsuka, R., Kanda, K., Yokoi, T., Kondo, H., Magatsuwa, M., Murakawa, K., Ishida, S., Ishibashi, T., Takahashi-Fujii, A., Tanase, T., Nagai, K., Kikuchi, H., Nakai, K., Isogai, T. and Sugano, S. Diversification of Transcriptional Modulation: Large-scale Identification and Characterization of Putative Alternative Promoters of Human Genes

JOURNAL PUBMED
Genome Res. 16 (1), 55-65 (2006)
16344560

COMMENT

Contact: Takao Isogai
FLJ Project (HRI Team)
Helix Research Institute
2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: flj-cdna@nifty.com
NEO human cDNA project (New Energy and Industrial Technology Developmental Organization, Japan); cDNA library construction: Helix Research Institute (HRI); 5'-end one pass sequencing: HRI, Research Association for Biotechnology (RAB) and Biotechnology Center, National Institute of Technology and Evaluation; 3'-end one pass sequencing: RAB.
Location/Qualifiers

FEATURES
source
1..546
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="FEBRA2002360"
/tissue_type="brain"
/dev_stage="fetal"
/clone_lib="FEBRA2"
/note="Vector: pME18SFL3"

ORIGIN

Query Match 87.6%; Score 18.4; DB 9; Length 546;
Best Local Similarity 95.0%; Pred. No. 7.5e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Cy 1 CAGGACAGCTGCGCTTTGA 20
|||||
Db 220 CAGGACTCAGTGGCTTTGA 239

RESULT 7

LOCUS DA722798 549 bp mRNA linear EST 15-NOV-2005
DEFINITION DA722798 NT2RI3 Homo sapiens cDNA clone NT2RI3003688 5', mRNA sequence.

ACCESSION

DA722798
DA722798.1 GI:82437260

KEYWORDS

EST.

SOURCE

Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

REFERENCE

AUTHORS

1 (bases 1 to 549)
Kimura, K., Wakamatsu, A., Suzuki, Y., Ota, T., Nishikawa, T., Yamashita, R., Yamamoto, J., Sekine, M., Teurillac, K., Wakaguri, H., Ishii, S., Sugiyama, T., Saito, K., Isono, Y., Irie, R., Kushida, N., Yoneyama, T., Otsuka, R., Kanda, K., Yokoi, T., Kondo, H., Magatsuwa, M., Murakawa, K., Ishida, S., Ishibashi, T., Takahashi-Fujii, A., Tanase, T., Nagai, K., Kikuchi, H., Nakai, K., Isogai, T. and Sugano, S. Diversification of Transcriptional Modulation: Large-scale Identification and Characterization of Putative Alternative Promoters of Human Genes
Genome Res. 16 (1), 55-65 (2006)

TITLE

JOURNAL
PUBMED
16344560

COMMENT

Contact: Takao Isogai
FLJ Project (HRI Team)
Helix Research Institute
2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986

Email: flj-cdna@nifty.com
NEO human cDNA project (New Energy and Industrial Technology Developmental Organization, Japan); cDNA library construction: Helix Research Institute (HRI); 5'-end one pass sequencing: HRI, Research Association for Biotechnology (RAB) and Biotechnology Center, National Institute of Technology and Evaluation; 3'-end one pass sequencing: RAB.
Location/Qualifiers

FEATURES

source
1..549
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="NT2RI3003688"
/cell_type="teratocarcinoma"
/cell_line="NT2"
/clone_lib="NT2RI3"
/note="Vector: pME18SFL3; majorly NT2 neuron; mRNA from NT2 neuronal precursor cells treated 2-weeks mitotic inhibitor after 5-weeks retinoic acid (RA) induction"

ORIGIN

Query Match 87.6%; Score 18.4; DB 9; Length 549;
Best Local Similarity 95.0%; Pred. No. 7.5e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Cy 1 CAGGACAGCTGCGCTTTGA 20
|||||
Db 220 CAGGACTCAGTGGCTTTGA 239

RESULT 8

LOCUS DA906377 583 bp mRNA linear EST 03-DEC-2005
DEFINITION DA906377 SKNMC2 Homo sapiens cDNA clone SKNMC2007458 5', mRNA sequence.

ACCESSION

DA906377
DA906377.1 GI:83037875

KEYWORDS

SOURCE

Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

REFERENCE

AUTHORS

1 (bases 1 to 583)
Kimura, K., Wakamatsu, A., Suzuki, Y., Ota, T., Nishikawa, T., Yamashita, R., Yamamoto, J., Sekine, M., Teurillac, K., Wakaguri, H., Ishii, S., Sugiyama, T., Saito, K., Isono, Y., Irie, R., Kushida, N., Yoneyama, T., Otsuka, R., Kanda, K., Yokoi, T., Kondo, H., Magatsuwa, M., Murakawa, K., Ishida, S., Ishibashi, T., Takahashi-Fujii, A., Tanase, T., Nagai, K., Kikuchi, H., Nakai, K., Isogai, T. and Sugano, S. Diversification of Transcriptional Modulation: Large-scale Identification and Characterization of Putative Alternative Promoters of Human Genes
Genome Res. 16 (1), 55-65 (2006)

TITLE

JOURNAL
PUBMED
16344560

COMMENT

Contact: Takao Isogai
FLJ Project (HRI Team)
Helix Research Institute
2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: flj-cdna@nifty.com
NEO human cDNA project (New Energy and Industrial Technology Developmental Organization, Japan); cDNA library construction: Helix Research Institute (HRI); 5'-end one pass sequencing: HRI, Research Association for Biotechnology (RAB) and Biotechnology Center, National Institute of Technology and Evaluation; 3'-end one pass sequencing: RAB.
Location/Qualifiers

FEATURES

source
1..583
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"


```

/molecule="genomic DNA"
/cultivar="B73"
/db_xref="taxon:4577"
/clone="ZMABF0080F02"
/lab_host="EP1100-T1"
/clone_1ib="ZMABF"
/note="Vector: pBp1FOS-5; Site_1: Eco72I"

ORIGIN

Query Match      87.6%; Score 18.4; DB 13; Length 654;
Best Local Similarity 95.0%; Pred. No. 7.6e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 2 AGGACACAGTGGCTTTGAC 21
    |||||
Db 579 AGGACACAGTGGCTTTGAC 598

RESULT 11
LOCUS C333727 660 bp mRNA linear EST 14-JUN-2005
DEFINITION C333727 Molgula tectiformis unpublished cDNA library Molgula
tectoformis cDNA clone mtbh004d21 3', mRNA sequence.
ACCESSION C333727
VERSION C333727.1 GI:67744326
KEYWORDS EST
SOURCE Molgula tectiformis
ORGANISM Molgula tectiformis
Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea;
Stolidobranchia; Molgulidae; Molgula.
1 (bases 1 to 660)
AUTHORS Gyoja, F., Satou, Y. and Satoh, N.
TITLE Expressed genes in Molgula tectiformis
JOURNAL Unpublished (2005)
COMMENT Contact: Tadasu Shin-i
National Institute of Genetics
Center For Genetic Resource Information
National Institute of Genetics
111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp
When you want to obtain this EST clone, please send an e-mail to
Nori Satoh (satoh@acidian.zool.kyoto-u.ac.jp) and its cc to Yutaka
Satou (yutaka@acidian.zool.kyoto-u.ac.jp).
Location/Qualifiers
1. 660
/organism="Molgula tectiformis"
/molecule="mRNA"
/db_xref="taxon:30286"
/clone="mtbh004d21"
/tissue_type="whole animal"
/dev_stage="embryo just before hatching"
/clone_1ib="Molgula tectiformis unpublished cDNA library"

FEATURES
SOURCE
1. 660
/organism="Molgula tectiformis"
/molecule="mRNA"
/db_xref="taxon:30286"
/clone="mtbh004d21"
/tissue_type="whole animal"
/dev_stage="embryo just before hatching"
/clone_1ib="Molgula tectiformis unpublished cDNA library"

ORIGIN

Query Match      87.6%; Score 18.4; DB 5; Length 660;
Best Local Similarity 95.0%; Pred. No. 7.7e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 CAGACACAGTGGCTTTGAC 20
    |||||
Db 43 CAGACACAGTGGCTTTGAC 62

RESULT 12
LOCUS CD353955 720 bp mRNA linear EST 15-JUL-2003
DEFINITION UT-M-GMO-gga-1-04-0-UI.r1 NIH_BMAP_GMO Mus musculus cDNA clone
IMAGE:30360123 5', mRNA sequence.
ACCESSION CD353955
VERSION CD353955.1 GI:31146456
KEYWORDS EST.
SOURCE Mus musculus (house mouse)

```

```

ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridea; Muridae; Murinae; Mus.
1 (bases 1 to 720)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cga@db-remail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
http://genome.uiowa.edu/distribution/mouse1.html
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)

Seq primer: PYX-5.
Location/Qualifiers
1. 720
/organism="Mus musculus"
/molecule="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:30360123"
/tissue_type="whole brain"
/dev_stage="1, 5 and 15 days newborn"
/lab_host="DH10B (T1 phage resistant)"
/clone_1ib="NIH BMAP GM0"
/note="Organ: Brain; Vector: pYX-Asc; Site 1: BclI;
Site 2: NotI; The library was constructed according
to Bonaldi, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured RNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with oligo-dT
primer containing a NotI site. Double strand cDNA was size
selected according to mRNA size fraction. Ligated with BclI
I adaptor, digested with NotI and then cloned
directionally into pYX-Asc vector. The library tag
sequence located between the NotI site and the polyA tail
is CCACTGAT. This library was created for the University
Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
Developing Mouse Nervous System', supported by National
Institute of Mental Health (NIMH), Hemin Chn, Ph.D.,
program coordinator."

ORIGIN

Query Match      87.6%; Score 18.4; DB 5; Length 720;
Best Local Similarity 95.0%; Pred. No. 7.8e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 2 AGGACACAGTGGCTTTGAC 21
    |||||
Db 495 AGGACACAGTGGCTTTGAC 514

RESULT 13
LOCUS AY406159 2148 bp DNA linear GSS 15-DEC-2003
DEFINITION Pan troglodytes CDH8 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION AY406159
VERSION AY406159.1 GI:39762133
KEYWORDS GSS.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Pan.
1 (bases 1 to 2148)
AUTHORS Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A.,
Todd, M.A., Tenenbaum, D.M., Civallo, D.R., Lu, P., Murphy, B.,
Ferreira, S., Wang, G., Zheng, X.H., White, T.J., Sinsky, J.D.,

```

TITLE Adams,M.D. and Cargill,M.
Infering nonneutral evolution from human-chimp-mouse orthologous
gene trios

JOURNAL Science 302 (5652), 1960-1963 (2003)

REFERENCE PUBMED 14671302

AUTHORS 2 (bases 1 to 2148)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civallo,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Shtinsky,J.J.,
Adams,M.D. and Cargill,M.

TITLE Direct Submission

JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA

COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.

FEATURES Location/Qualifiers

1..2148

source /organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
<1..>2148
/gene="CDH8"
/locus_tag="HCM2464"

ORIGIN

Query Match 87.6%; Score 18.4; DB 14; Length 2148;
Best Local Similarity 95.0%; Pred. No. 9.1e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 AGGACACAGTGGCTTTGAC 21
|||||
Db 1772 AGGACACAGGCGTTTGAC 1791

RESULT 14

AY406158 2400 bp DNA linear GSS 15-DEC-2003

LOCUS Homo sapiens CDH8 gene, VIRTUAL TRANSCRIPT, partial sequence.

DEFINITION Genomic survey sequence.

ACCESSION AY406158

VERSION AY406158.1 GI:39762132

KEYWORDS GSS.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 2400)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civallo,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Shtinsky,J.J.,
Adams,M.D. and Cargill,M.

AUTHORS Adams,M.D. and Cargill,M.

TITLE Infering nonneutral evolution from human-chimp-mouse orthologous
gene trios

JOURNAL Science 302 (5652), 1960-1963 (2003)

REFERENCE PUBMED 14671302

AUTHORS 2 (bases 1 to 2400)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civallo,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Shtinsky,J.J.,
Adams,M.D. and Cargill,M.

TITLE Direct Submission

JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA

COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.

FEATURES Location/Qualifiers

1..2400

source /organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
<1..>2400
/gene="CDH8"

ORIGIN

ORIGIN /locus_tag="HCM2464"

Query Match 87.6%; Score 18.4; DB 14; Length 2400;
Best Local Similarity 95.0%; Pred. No. 9.2e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 AGGACACAGTGGCTTTGAC 21
|||||
Db 2024 AGGACACAGGCGTTTGAC 2043

RESULT 15

AY406160 2400 bp DNA linear GSS 15-DEC-2003

LOCUS Mus musculus CDH8 gene, VIRTUAL TRANSCRIPT, partial sequence.

DEFINITION Genomic survey sequence.

ACCESSION AY406160

VERSION AY406160.1 GI:39762134

KEYWORDS GSS.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 2400)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civallo,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Shtinsky,J.J.,
Adams,M.D. and Cargill,M.

AUTHORS Adams,M.D. and Cargill,M.

TITLE Infering nonneutral evolution from human-chimp-mouse orthologous
gene trios

JOURNAL Science 302 (5652), 1960-1963 (2003)

REFERENCE PUBMED 14671302

AUTHORS 2 (bases 1 to 2400)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civallo,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Shtinsky,J.J.,
Adams,M.D. and Cargill,M.

TITLE Direct Submission

JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA

COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.

FEATURES Location/Qualifiers

1..2400

source /organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
<1..>2400
/gene="CDH8"
/locus_tag="HCM2464"

ORIGIN

Query Match 87.6%; Score 18.4; DB 14; Length 2400;
Best Local Similarity 95.0%; Pred. No. 9.2e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 AGGACACAGTGGCTTTGAC 21
|||||
Db 2024 AGGACACAGGCGTTTGAC 2043

RESULT 16

AK139334 2711 bp mRNA linear HTC 21-SEP-2005

LOCUS Mus musculus 10 days neonate cortex cDNA, RIKEN full-length
enriched library, clone:A830013124 product:cadherin 8, full insert
sequence.

ACCESSION AK139334

VERSION AK139334.1 GI:74228151

KEYWORDS HTC; CAP trapper.

SOURCE Mus musculus (house mouse)

[illegible]

TITLE
JOURNAL
 Direct Submission
 Submitted (30-MAR-2004) Yoshihide Hayashizaki, The Institute of
 Physical and Chemical Research (RIKEN), Laboratory for Genome
 Exploration and Chemical Research Group, RIKEN Genomic Sciences Center (GSC),
 RIKEN Yokohama Institute; 1-7-22 Suenhiro-cho, Tsunumi-ku, Yokohama,
 Kanagawa, 230-0045, Japan (E-mail:genome-res@sc.riken.jp,
 URL:http://genome.gsc.riken.jp/, Tel:81-45-503-9222,
 Fax:81-45-503-9216)

COMMENT
 cDNA library was prepared and sequenced in Mouse Genome
 Encyclopedia Project of Genome Exploration Research Group in Riken
 Genomic Sciences Center and Genome Science Laboratory in Riken.
 Division of Experimental Animal Research in Riken contributed to
 prepare mouse tissues.
 Please visit our web site for further details.
 URL:http://genome.gsc.riken.jp/
 URL:http://fantom.gsc.riken.jp/
 Location/Qualifiers

FEATURES
source
 1..2711
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="PANTOM,DB:AB30013124"
 /db_xref="taxon:10090"
 /clone="AB30013124"
 /tissue_type="cortex"
 /clone_lib="RIKEN full-length enriched mouse cDNA library"
 /dev_stage="10 days neonate"
 547..2682
 /note="unnamed protein product; cadherin 8 (MGD|MGJ:107434
 GB|AK083092, evidence: BLASTN, 99%, match=2635)
 putative"
 /codon_start=1
 /protein_id="BAE23961.1"
 /db_xref="GI:74228152"
 /translation="MPERLATMDLWPTLLITLTLPSCTYTAAMNANHLYTGSPL
 ELRSSEDMRLTSLKRGVWVNFVLESGPEFLLVGLATDLDPSKIKITLISG
 DGAGTIFQINDITGDIAIKRLDREKSEYTLTAQAVFETNKPLPSSBPIIKVODI
 NDNAPEFLQNPYHATVPMSILGTSVNTATADDPYGSAXLVSLISQDPFSI
 EPEFALIKTALPNMDRAKEEYLVVIOAKMGHSGSGSTTLVTLVDNDPFS
 AOSLVHSPVPDVVGTALIGRYKANDDIGNAOSVYDIIIGDGLALEITSDAOD
 GVIRLRKLPDFTKSTYTLKVRANIHIDSSRSSGPKPDATYKIVVEDADPEVRS
 SPYLVLEHNAALNSVIGQYVATPDPTSSPIRSDIRHTDLESQFINADGKITL
 ATPIDRELVSNNHITITATEIRNHSQISRPVALKVLVDNNAPEFASVAFICENG
 KPGQVITQVSAMDQDPKQGHFLYSLLPEWNNPNPIKKNEDMSLSILAKNGFNR
 QKQEVYLPVIYSDGNPLSTSTLTITRVCCSNDGVQSCNVAAYVLPIGLSMGAL

Query Match
 Best Local Similarity 87.6%; Score 18.4; DB 6; Length 2711;
 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OR
 2 AGGACACAGTGGCTTTGAC 21
 |||||||
Db
 2570 AGGACACAGAGGCTTTGAC 2589

RESULT 17
AK044046 3161 bp mRNA linear HTC 02-SEP-2005
LOCUS
DEFINITION
 Mus musculus 10 days neonate cortex cDNA, RIKEN full-length
 enriched library, clone:AB3008313 product:cadherin 8, full insert
 sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
 Mus musculus (house mouse)
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
AUTHORS
TITLE
 Carninci, P. and Hayashizaki, Y.
 High-efficiency full-length cDNA cloning

JOURNAL
PUBMED
 10349636
REFERENCE
AUTHORS
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
 Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL
PUBMED
 11042159
REFERENCE
AUTHORS
 Shibata, K., Itoh, M., Aizawa, K., Nagao, S., Sasaki, N., Carninci, P.,
 Kono, H., Akiyama, U., Nishi, K., Katsunuma, T., Tashiro, H., Itoh, M.,
 Suni, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
 Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
 Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Matshiki, M.,
 Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
 Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
 RIKEN integrated sequence analysis (RISA) system-384-format
 sequencing pipeline with 384 multicapillary sequencer
JOURNAL
PUBMED
 11076861
REFERENCE
AUTHORS
 The RIKEN Genome Exploration Research Group Phase II Team and the
 PANTOM Consortium.
TITLE
 Functional annotation of a full-length mouse cDNA collection
JOURNAL
PUBMED
 Nature 409, 685-690 (2001)
REFERENCE
AUTHORS
 The PANTOM Consortium, the RIKEN Genome Exploration Research Group
 Phase I and II Team.
TITLE
 Analysis of the mouse transcriptome based on functional annotation
 of 60,770 full-length cDNAs
JOURNAL
PUBMED
 Nature 420, 563-573 (2002)
REFERENCE
AUTHORS
 RIKEN Genome Exploration Research Group, Genome Science Group
 (Genome Network Core Team) and the PANTOM Consortium.
TITLE
 Antisense Transcription in the Mammalian Transcriptome
JOURNAL
PUBMED
 Science 309, 1564-1566 (2005)
REFERENCE
AUTHORS
 The PANTOM Consortium, Riken Genome Exploration Research Group and
 Genome Science Group (Genome Network Project Core Group).
TITLE
 The Transcriptional Landscape of the Mammalian Genome
JOURNAL
PUBMED
 Science 309, 1559-1563 (2005)
REFERENCE
AUTHORS
 8 (bases 1 to 3161)
 Adachi, J., Aizawa, K., Akimura, T., Arikawa, T., Bono, H., Carninci, P.,
 Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
 Hayashida, K., Hayatsu, N., Hiramoto, K., Hirooka, T., Hirozane, T.,
 Horii, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
 Kato, H., Kawai, J., Kojima, Y., Kondo, S., Kono, H., Kouda, M.,
 Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
 Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Ohsato, N.,
 Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, R., Sakazume, N.,
 Sano, H., Sasaki, D., Shibata, K., Shingawa, A., Shiraki, T.,
 Sogabe, Y., Tagami, M., Tagawa, A., Takeda, F., Takahashi, A., S.,
 Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
 Muramatsu, M. and Hayashizaki, Y.
TITLE
 Direct Submission
JOURNAL
 Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
 Physical and Chemical Research (RIKEN), Laboratory for Genome
 Exploration and Chemical Research (RIKEN), Laboratory for Genome
 Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
 RIKEN Yokohama Institute; 1-7-22 Suenhiro-cho, Tsunumi-ku, Yokohama,
 Kanagawa, 230-0045, Japan (E-mail:genome-res@sc.riken.jp,
 URL:http://genome.gsc.riken.jp/, Tel:81-45-503-9222,
 Fax:81-45-503-9216)

COMMENT
 cDNA library was prepared and sequenced in Mouse Genome
 Encyclopedia Project of Genome Exploration Research Group in Riken
 Genomic Sciences Center and Genome Science Laboratory in Riken.
 Division of Experimental Animal Research in Riken contributed to
 prepare mouse tissues.
 Please visit our web site for further details.
 URL:http://genome.gsc.riken.jp/
 URL:http://fantom.gsc.riken.jp/
 Location/Qualifiers

FEATURES
source
 1..3161

```

/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="PANTOM_DB:A830083P13"
/db_xref="taxon:10090"
/clone="A830083P13"
/issue_type="cortex"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_offset="10 days neonate"
549..2813
/note="unnamed protein product; cadherin 8 (MGD|MG1:107434
GB|NM_007667, evidence: BLASTN, 99%, match=2357)
putative"
/codon_start=1
/protein_id="BAC31751.1"
/db_xref="gi:2636132"
/translation="MPERLAKTTMDLTPLIIITWITPSCVYTA PMNQAHVLTGSGPL
ELSRQSDMRILSRKGMWNOMFVLEEFSGPPIVGRILHTLDGSKIKYILSG
DGAQTIPOINDITGDIIHAKLDREERAEYLTQAVDPETNKLPEPSEKIKYDII
NDNAPELNGPYHATVPMSGLTGSYNTVATDADVDYNGSAKYVSIIEGQYESI
EPETAIITKTLALPNMDREKERYLYVITQAKMGHSGGLSGTTITVTITVDNPPKPF
AQSLYHPSVPEEDVVLGTAIGRVKANDDIDENASSTDIDEGDTALPEITSDAQAD
GVIRLRKPLPFEKTSYTLKVEANAIHIDPFSSRGPKDTAYKIVEBADPEPVS
SPTYLAEHENAALNSYGVGTADDPITSEPIFSDIHTLERQPINADSKITLT
ATPLDRELHVMNITITITATELRHNSOISRVVALKYVDNDNAEPASBEVALFCENG
KRGVITVYSAMDODPRNGRPFYLSILPEVNNPNTIKGNBNSIILAKNRPGR
OKQVYLTPLVITSDSGNPUSYSTLTITRVCGSNDVQSCNVEAYVLPIGLSMGL
IATACITLILVIVLPTTLRRHNBLIKDDVRENIIRYDDEGGSEDTFAFDI
ATLQNDPQINGFLPRKDIKPLQPMPLPKRLPRNTRRECLKKGTMDPSDFKEKIST
ODFRKSEVTKGTC"
3132..3137
/note="putative"
3161
/note="putative"
polya_site
polya_signal
ORIGIN
Query Match      87.6%; Score 18.4; DB 6; Length 3161;
Best Local Similarity 95.0%; Pred. No. 9.6e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Oy      2 AGGACACAGTGCGCTTTGAC 21
      ||||| ||||| |||||
Db      2572 AGGACACAGAGGCTTTTGAC 2591

RESULT 18
AC083092 LOCUS      3592 bp      mRNA      linear      HTC 02-SEP-2005
DEFINITION Mus musculus adult male hippocampus cDNA, RIKEN full-length
enriched library, clone: C63002D14 product: cadherin 8, full insert
sequence.
AC083092 VERSION
KEYWORDS Mus musculus (house mouse)
SOURCE HTC; CAP trapper.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridea; Muridae; Murinae; Mus.
REFERENCE 1 Carninci, P. and Hayashizaki, Y.
AUTHORS Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
PUBMED 10349636
REFERENCE 2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
AUTHORS Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
PUBMED 11042159
REFERENCE 3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
AUTHORS

```

TITLE	RIKEN Genome Exploration Research Group Phase II Team and the PANTOM Consortium.
JOURNAL	Nature 409, 685-690 (2001)
REFERENCE	
AUTHORS	
TITLE	The PANTOM Consortium, the RIKEN Genome Exploration Research Group Phase I and II Team.
JOURNAL	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
REFERENCE	Nature 420, 563-573 (2002)
AUTHORS	
TITLE	RIKEN Genome Exploration Research Group, Genome Science Group (Genome Network Core Team) and the PANTOM Consortium.
JOURNAL	Antisense Transcription in the Mammalian Transcriptome
REFERENCE	Science 309, 1564-1566 (2005)
AUTHORS	
TITLE	The PANTOM Consortium, Riken Genome Exploration Research Group and Genome Science Group (Genome Network Project Core Group).
JOURNAL	The Transcriptional Landscape of the Mammalian Genome
REFERENCE	Science 309, 1559-1563 (2005)
AUTHORS	
TITLE	8 (bases 1 to 3592)
JOURNAL	Adachi, S., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., . . . Furuta, S., Furuta, N., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, P., Imtani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Komori, H., Kondo, M., Koyama, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishii, K., Nomura, K., Numazaki, R., Ohno, M., Ohnishi, N., Ozaki, Y., Saito, R., Saitoh, K., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Saeki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, Y., Tanaka, T., Tomaru, A., Toyota, T., Yasunishi, A., Muramatsu, M., and Hayashizaki, Y.
REFERENCE	Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa, 230-0045, Japan (E-mail:genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
AUTHORS	
TITLE	CNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
JOURNAL	Please visit our web site for further details.
REFERENCE	URL: http://genome.gsc.riken.jp/
AUTHORS	
TITLE	location/Qualifiers
JOURNAL	1. 3592
REFERENCE	/organism="Mus musculus"
AUTHORS	/mol_type="mRNA"
TITLE	/strain="C57BL/6J"
JOURNAL	/db_xref="PANTOM DB: C630002D14"
REFERENCE	/db_xref="taxon:10090"
AUTHORS	/clone="C630002D14"
TITLE	/sex="male"
JOURNAL	/tissue type="hippocampus"
REFERENCE	/clone_id="RIKEN full-length enriched mouse cDNA library"
AUTHORS	/dev_stage="adult"
TITLE	558. 2708

```

/note="unnamed protein product; cadherin 8 [MGD|MG1:107434
GB|NM_007667, evidence: BLASTN, 99%, match=2120)
putative"
/codon_start=1
/db_xref="BAC39758.1"
/translation="MPERLAFTLMDLMTPLIILMTLPSCVYTAPMNOAHVLTGSP
ELRSQSEDMRLISRGKGMVNMQVLEEFSGTPEILVGRILATDLDPSGKIKYLSG
DAGTIFQINDITGDIIHAKRLDREKAEYTLTQAVDFENKPLEPSPFIKQDI
NDNAPEFLNGPYHATVPKMSILGTSVTNTVATDADDPYGNASAKLVYSILSGQPFESI
EPETAILKTALPNMDREKREYLVVIOAKDNGSGSGGSGTTLVTTLTVNDNPKF
AQSILHPSVPEDVVLGTAIGRVKANDDIGNAOSSTIITDGTALFEITSDAQAD
GVILRKPLDETKKSYTLKYVAANIHDIPRSPSGPKDTATVIVIVEDADPEVFS
SPYTLAEVHENAALNSVIGQVATDPDITSPSPRPSIDRHLDLEKQFINADGKITL
ATPLDRELSEVHNITITATETIRNHSQISRPVAVLVDNNAPEFASVEAFCENG
KPGQVITQVSAAMDQDPRNGHFLYSLPEVNNPNPTIKKQEDNSILAKNGFNR
OKQRYVILPIVSDSGNPLSTSTLTIRVCGCSDNVDGSCNVEAYLVLPGLSGMAL
IATLACTILLVIVLFTVTLRHRKREPLITIKODEVRENITIRYDEGGEDTEAFDI
ATLQNPDSINGPLPRWGLPTORWBERNAIQGCV"

ORIGIN
Query Match      87.6%  Score 18.4;  DB 6;  Length 3592;
Best Local Similarity 95.0%  Pred. No. 9.8e+02;
Matches 19;  Conservative 0;  Mismatches 1;  Indels 0;  Gaps 0;

Oy      2 AGGACACAGTGGCTTTGAC 21
        |||||
Db      2581 AGGACACAGAGGCTTTGAC 2600

RESULT 19
AK086711      3793 bp  mRNA  linear  HTC 02-SEP-2005
LOCUS      Mus musculus 15 days embryo head cDNA, RIKEN full-length enriched
DEFINITION library, clone:DB30046N17 product:cadherin 8, full insert sequence.
ACCESSION  AK086711
VERSION     AK086711.1 GI:26352174
KEYWORDS    HTC; CAP trapper.
SOURCE      Mus musculus (house mouse)
ORGANISM    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muridae; Murinae; Mus.

REFERENCE   1
  AUTHORS   Carninci,P. and Hayashizaki,Y.
  TITLE     High-efficiency full-length cDNA cloning
  JOURNAL   Meth. Enzymol. 303, 19-44 (1999)
  PUBMED   10349636
  REFERENCE
  AUTHORS   2
  TITLE     Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
            Itoh,M., Kono,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
            Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new genes
            JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
            PUBMED 11042159
            REFERENCE
            AUTHORS   3
            TITLE     Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
            Kono,H., Akiyama,J., Nishi,K., Kitsuana,T., Tashiro,H., Itoh,M.,
            Suni,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
            Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
            Fujiwara,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,S., Watanabe,M.,
            Okazaki,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura,S., Kawai,J.,
            Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
            RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer
            JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
            PUBMED 11076861
            REFERENCE
            AUTHORS   4
            TITLE     The RIKEN Genome Exploration Research Group Phase II Team and the
            FANTOM Consortium.
            Functional annotation of a full-length mouse cDNA collection
            JOURNAL Nature 403, 685-690 (2001)
            REFERENCE
            5

```

```

AUTHORS      The FANTOM Consortium, the RIKEN Genome Exploration Research Group
TITLE        Phase I and II Team.
JOURNAL      Analysis of the mouse transcriptome based on functional annotation
            of 60,770 full-length cDNAs
            Nature 420, 563-573 (2002)
REFERENCE    6
  AUTHORS    RIKEN Genome Exploration Research Group, Genome Science Group
            (Genome Network Core Team) and the FANTOM Consortium.
            Antisense Transcription in the Mammalian Transcriptome
            Science 309, 1564-1566 (2005)
REFERENCE    7
  AUTHORS    The FANTOM Consortium, Riken Genome Exploration Research Group and
            Genome Science Group (Genome Network Project Core Group).
            The transcriptional landscape of the Mammalian Genome
            Science 309, 1559-1563 (2005)
            8 (bases 1 to 3793)
            Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,
            Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,M.,
            Hayashida,K., Hayatsu,N., Hiramoto,K., Hirooka,T., Hirozane,T.,
            Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T.,
            Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Kono,H., Kouda,M.,
            Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
            Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N.,
            Okazaki,Y., Saito,K., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N.,
            Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
            Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,
            Takeda,Y., Tanaka,T., Tomaru,A., Toyu,T., Yasunishi,A.,
            Muramatsu,M. and Hayashizaki,Y.
            Direct Submission
            Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of
            Physical and Chemical Research (RIKEN), Laboratory for Genome
            Exploration and Research Group, RIKEN Genomic Sciences Center (GSC),
            RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
            Kanagawa, 230-0045, Japan (E-mail:genome-res@gsc.riken.jp,
            URL:http://genome.gsc.riken.jp/, Tel:81-45-503-9222,
            Fax:81-45-503-9216)
            cDNA library was prepared and sequenced in Mouse Genome
            Encyclopedia Project of Genome Exploration Research Group in Riken
            Genomic Sciences Center and Genome Science Laboratory in Riken
            Division of Experimental Animal Research in Riken contributed to
            prepare mouse tissues.
            Please visit our web site for further details.
            URL:http://genome.gsc.riken.jp/.
            URL:http://fantom.gsc.riken.jp/.

FEATURES
  source      location/Qualifiers
              1..3793
              /organism="Mus musculus"
              /mol_type="mRNA"
              /strain="C57BL/6J"
              /db_xref="FANTOM_DB:DB30046N17"
              /db_xref="taxon:10090"
              /clone="DB30046N17"
              /tissue_type="head"
              /clone_lib="RIKEN full-length enriched mouse cDNA library"
              /dev_stage="15 days embryo"
              560..2710
              /note="unnamed protein product; cadherin 8
              [MGD|MG1:107434, evidence: Direct assignment)
              putative"
              /codon_start=1
              /protein_id="BAC39724.1"
              /db_xref="GI:26352175"
              /translation="MPERLAFTLMDLMTPLIILMTLPSCVYTAPMNOAHVLTGSP
              ELRSQSEDMRLISRGKGMVNMQVLEEFSGTPEILVGRILATDLDPSGKIKYLSG
              DAGTIFQINDITGDIIHAKRLDREKAEYTLTQAVDFENKPLEPSPFIKQDI
              NDNAPEFLNGPYHATVPKMSILGTSVTNTVATDADDPYGNASAKLVYSILSGQPFESI
              EPETAILKTALPNMDREKREYLVVIOAKDNGSGSGGSGTTLVTTLTVNDNPKF
              AQSILHPSVPEDVVLGTAIGRVKANDDIGNAOSSTIITDGTALFEITSDAQAD
              GVILRKPLDETKKSYTLKYVAANIHDIPRSPSGPKDTATVIVIVEDADPEVFS
              SPYTLAEVHENAALNSVIGQVATDPDITSPSPRPSIDRHLDLEKQFINADGKITL
              ATPLDRELSEVHNITITATETIRNHSQISRPVAVLVDNNAPEFASVEAFCENG
              KPGQVITQVSAAMDQDPRNGHFLYSLPEVNNPNPTIKKQEDNSILAKNGFNR
              OKQRYVILPIVSDSGNPLSTSTLTIRVCGCSDNVDGSCNVEAYLVLPGLSGMAL

```

ORIGIN
 IAILACIILLIVIVLFTLRRHKNRPLIKODEDLRNIIRYDEGGSEBETBAFDI
 ATLQNPDGINGLPRWGLPTQWEERNALQTCGV"

Query Match 87.6%; Score 18.4; DB 6; Length 3793;
 Best Local Similarity 95.0%; Pred. No. 9, 8e+02;
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AGGACACAGTGGCTTTGAC 21
 |||||
 DB 2583 AGGACACAGAGGCTTTGAC 2602

RESULT 20
 BG022939/c 337 bp mRNA linear EST 24-JAN-2001

LOCUS da81d07.x1 Cho Li treated gastrula xenopus laevis cDNA clone
 DEFINITION IMAGE:4083876 3' similar to TR:Q92688 Q92688 PAFPIZA PROTEIN ;
 mRNA sequence.

ACCESSION BG022939.1 GI:12479018
 VERSION EST.
 KEYWORDS Xenopus laevis (African clawed frog)
 ORGANISM Xenopus laevis

REFERENCE 1 (bases 1 to 337)
 AUTHORS Clifton, S., Johnson, S.L., Blumberg, B., Song, J., Hillier, L.,
 Pape, D., Martin, J., Wylie, T., Underwood, K., Theising, B., Bowers, Y.,
 Person, B., Gibbons, M., Harvey, N., Rittner, E., Jackson, Y., McCann, R.,
 Waterston, R. and Wilson, R.
 Washu Xenopus EST project, 1999
 Unpublished (1999)

TITLE Contact: Sandy Clifton, Ph.D.
 JOURNAL Washu Xenopus EST project, 1999
 COMMENT Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 Library constructed by C. Hashimoto, Ph.D. in the laboratory of K.
 Cho, Ph.D. DNA Sequencing by: Washington University Genome
 Sequencing Center

FEATURES
 source
 1..337
 /organism="Xenopus laevis"
 /mol_type="mRNA"
 /db_xref="taxon:8355"
 /clone="IMAGE:4083876"
 /tissue_type="gastrula, li-treated"
 /lab_host="DH10B"
 /clone_lib="Cho Li treated gastrula"
 /note="Vector: pBluescript KS+; Site 1: NotI; Site 2:
 SCOR1; 1st strand was primed with a Not I - oligo (dT)
 primer, double-stranded cDNA was cloned into the Not I and
 Eco RI sites of pBluescript KS+ library was constructed
 by C. Hashimoto, Ph.D., in the laboratory of K. Cho, Ph.D.
 (Department of Developmental and Cell Biology, University
 of California, Irvine)."

ORIGIN

Query Match 84.8%; Score 17.8; DB 2; Length 337;
 Best Local Similarity 90.5%; Pred. No. 1, 4e+03;
 Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 1 CAGGACACAGTGGCTTTGAC 21
 |||||

DB 25 CAGGACAAAGAGGCTTTGAC 5

RESULT 21
 BF771220/c 357 bp mRNA linear EST 12-JAN-2001

LOCUS IL5-IT0027-271100-285-d06 IT0027 Homo sapiens cDNA, mRNA sequence.
 DEFINITION BF771220
 ACCESION BF771220.1 GI:12119120
 VERSION EST.
 KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homiidae; Homo.

REFERENCE 1 (bases 1 to 357)
 AUTHORS Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
 Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
 Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H.,
 Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V.,
 O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
 Simpson, A.J.

TITLE Shotgun sequencing of the human transcriptome with ORF expressed
 sequence tags
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 PUBMED 10737800
 COMMENT Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome
 Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?cl=IL5&t2=IL5-IT0027-
 271100-285-d06&t3=2000-11-27&t4=1)
 Seq primer: puc 18 forward
 High quality sequence stop: 302.

FEATURES
 source
 1..357
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /dev_stage="Adult"
 /clone_lib="IT0027"

FEATURES
 source
 1..357
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /dev_stage="Adult"
 /clone_lib="IT0027"
 /note="Organ: epid tumor; Vector: puc18; Site 1: SmaI;
 Site 2: SmaI; A mini-library was made by cloning products
 derived from ORSTES PCR (U.S. Letters Patent application
 No. 196,716 - Ludwig Institute for Cancer Research)
 profiles into the puc 18 vector. Reverse transcription of
 tissue mRNA and cDNA amplification were performed under
 low stringency conditions."

ORIGIN

Query Match 84.8%; Score 17.8; DB 2; Length 357;
 Best Local Similarity 90.5%; Pred. No. 1, 4e+03;
 Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 1 CAGGACACAGTGGCTTTGAC 21
 |||||
 DB 56 CAGGACCAAGTGGCTTTGAC 36

RESULT 22

CR555978 377 bp mRNA linear EST 12-JUL-2004
 LOCUS DKR2p4691192.r1 469 (synonym: pkid1) Pongo pygmaeus cDNA clone
 DEFINITION DKR2p4691192 5', mRNA sequence.
 ACCESION CR555978
 VERSION CR555978.1 GI:50249580
 KEYWORDS EST.

SOURCE
ORGANISM
Pongo pygmaeus (orangutan)

REFERENCE
AUTHORS
1 (bases 1 to 377)
Boeckler, H., Boecher, M., Brandt, P., Mewes, H.W., Weill, B., Amd, C., Oesinger, A., Fob, G., Han, M., and Wiemann, S.
Pongo pygmaeus mRNA (Boeckler, H., Boecher, M., Brandt, P., et al.)
Unpublished (2004)
Contact: MIPS

JOURNAL
COMMENT
MIPS
Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert Clone from S. Wiemann, Molecular Analysis, German Cancer Research Center (DKFZ), Email s.wiemann@dkfz-heidelberg.de; sequenced by GBF (National Research Centre for Biotechnology Ltd., Braunschweig/Germany) within the cDNA sequencing consortium of the German Genome Project. This clone (DKFZp4691192) is available at the RZPD in Berlin. Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further information about the clone and the sequencing project is available at <http://mips.gsf.de/projects/cdna/>.
Location/Qualifiers

FEATURES
source
1..377
/organism="Pongo pygmaeus"
/mol_type="mRNA"
/db_xref="taxon:9600"
/clone="DKFZp4691192"
/tissue_type="kidney"
/dev_stage="adult"
/lab_host="DH10B"
/clone_id="469 (synonym: pkidi)"
/note="Vector: pSPORT1_sfi; site_1: sf1a; site_2: sf1b"

ORIGIN
Query Match 84.8%; Score 17.8; DB 8; Length 377;
Best Local Similarity 90.5%; Pred. No. 1.4e+03;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
1 CAGGACACAGTGGCTTTGAC 21
54 CAGGACACAGTGGCTTTGAC 74

RESULT 23
BY467982 501 bp mRNA linear EST 31-DEC-2002
LOCUS BY467982 RIKEN full-length enriched, melanocyte Mus musculus cDNA
DEFINITION BY467982 G270040C21 3', mRNA sequence.
ACCESSION BY467982
VERSION BY467982.1 GI:26802361
KEYWORDS BSR.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Muridae; Murinae; Mus.
1 (bases 1 to 501)
Okazaki, Y., Furuno, M., Kaenkawa, T., Adachi, J., Bono, H., Kondo, S., Nishio, I., Otsu, N., Saito, R., Suzuki, H., Yamashita, T., Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schombach, C., Gojobori, T., Baldarelli, R., Hill, D.P., Bult, C., Hume, D.A., Quackenbush, J., Schriml, L.M., Kanapin, A., Matcude, H., Baralov, S., Beisel, K.W., Blake, J.A., Bradt, D., Brusic, V., Chothia, C., Corbani, L.E., Cousine, S., Della, B., Dragani, T.A., Fletcher, C.F., Forrest, A., Frazer, K.S., Gasteiger, T., Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S., Guetlich, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A., Kawai, H., Kawasawa, Y., Kedzierski, R.M., King, B.L., Kongave, A., Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R., Maltas, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T., Numata, K., Okido, T., Pavan, W.J., Petca, G., Peeble, G.,

Petrovsky, N., Pillai, R., Pontius, J.V., Qi, D., Ramchandran, S., Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring, B.Z., Ringwald, M., Sandelin, A., Schneider, C., Sempile, C.A., Setou, M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M., Verardo, R., Wagner, L., Wahlested, C., Wang, Y., Warande, Y., Wells, C., Wilming, L.G., Wynshaw-Boris, A., Yang, I., Yang, L., Yuan, Z., Zavoian, M., Zhu, Y., Zimmer, A., Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Saeki, D., Shibata, K., Shingawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S., Rogers, J., Birney, E., and Hayashizaki, Y.
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
12466851
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Saitama-cho, Tsukuba, Ibaraki, 305-8575, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gs.c.riken.jp, URL: <http://genome.gsc.riken.jp/>
Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hirozane, T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Konno, H., Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numata, K., Ohno, M., Sakai, K., Sakazume, N., Saeki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M., Waki, K., Watanabe, A., Muramatsu, M., and Hayashizaki, Y.
Direct Submision
Computational Analysis of Full-length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)
RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)
Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Cells were provided by Drs. William J Pavan, Stacie Loftus, and Denise Larson (Division of Intramural Research Genetic Disease Research Branch National Human Genome Research Institute, National Institutes of Health (NIH) Building: 49, Room 4A82 49 Convent Drive MSC 4472 Bethesda, Maryland U.S.A) whose assistance we gratefully acknowledge
Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.
Location/Qualifiers
1..501
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="G270040C21"
/tissue_type="melanocyte"
/clone_id="RIKEN full-length enriched, melanocyte"

FEATURES
source
1..501
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="G270040C21"
/tissue_type="melanocyte"
/clone_id="RIKEN full-length enriched, melanocyte"

ORIGIN
Query Match 84.8%; Score 17.8; DB 4; Length 501;
Best Local Similarity 90.5%; Pred. No. 1.5e+03;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
1 CAGGACACAGTGGCTTTGAC 21
186 CAGGACACAGTGGCTTTGAC 206

RESULT 24
 CR429603/c 507 bp mRNA linear EST 17-JUN-2004
 LOCUS CR429603 XGC-tailbud Xenopus tropicalis cDNA clone TTB075007 5',
 DEFINITION mRNA sequence.
 CR429603
 ACCESSION CR429603.1 GI:48923012
 VERSION EST.
 KEYWORDS Xenopus tropicalis (western clawed frog)
 SOURCE Xenopus tropicalis
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
 Xenopodinae; Xenopus; Silurana.
 REFERENCE 1 (bases 1 to 507)
 AUTHORS Croning,M.D.R., Ashurst,J.L., Taylor,R., Garrett,N. and Rogers,J.
 TITLE Sanger Xenopus tropicalis EST project 2001 (2004)
 JOURNAL Unpublished (2004)
 COMMENT Contact: Croning MDR
 Sanger Institute
 Hinxton, Cambridgeshire, CB10 1SA, UK
 Email: trop@sanger.ac.uk
 Sanger Xenopus tropicalis EST project 2001
 TROPICALIS_SEQUENCE_ID: TTB075007.plksp6
 This sequence is from a Xenopus Gene Collection (XGC) library
 constructed by Nigel Garrett.
 Seq primer: SP6.
 ORIGIN
 location/Qualifiers
 1..507
 /organism="Xenopus tropicalis"
 /mol_type="mRNA"
 /db_xref="taxon:8364"
 /clone="TTB075007"
 /dev_stage="tailbud (stage 28-30)"
 /lab_host="Escherichia coli DH10B."
 /clone_lib="XGC-tailbud"
 /note="Vector: pCS107; Site_1: EcoRI; Site_2: NotI; cDNA
 was oligo dt primed from 5ug of poly A+ RNA from neurula.
 EcoRI-NotI cut cDNA was then ligated into pCS107 with
 EcoRI at the 5' end and NotI at the 3' end."
 ORIGIN

Query Match 84.8%; Score 17.8; DB 8; Length 507;
 Best Local Similarity 90.5%; Pred. No. 1.5e+03;
 Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 CAGGACACAGTGGCTTTGAC 21
 |||||
 Db 340 CAGGACACAGTGGCTTTGAC 320

RESULT 25
 AL785427/c 594 bp mRNA linear EST 13-NOV-2003
 LOCUS AL785427 XGC-neurula Xenopus tropicalis cDNA clone TNeu072e23 5',
 DEFINITION mRNA sequence.
 AL785427
 ACCESSION AL785427.2 GI:38308397
 VERSION EST.
 KEYWORDS Xenopus tropicalis (western clawed frog)
 SOURCE Xenopus tropicalis
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
 Xenopodinae; Xenopus; Silurana.
 REFERENCE 1 (bases 1 to 594)
 AUTHORS Croning,M.D.R., Ashurst,J.L., Taylor,R., Zorn,A.M. and Rogers,J.
 TITLE Sanger Xenopus tropicalis EST project 2001 (11_2003)
 JOURNAL Unpublished (2003)
 COMMENT On Jun 25, 2002 this sequence version replaced gi:21571131.
 Contact: Taylor R
 Sanger Institute
 Hinxton, Cambridgeshire, CB10 1SA, UK

Email: trop@sanger.ac.uk
 This sequence is from a Xenopus Gene Collection (XGC) library
 constructed by Aaron M. Zorn.
 cDNA was oligo dt primed from 5ug of poly A+ RNA from neurula.
 EcoRI-NotI cut cDNA was then ligated into pCS107 with EcoRI at the
 5' end and NotI at the 3' end.
 Vector: pCS107; Site_1: EcoRI; Site_2: NotI
 Host: Escherichia coli DH10B
 Sanger Xenopus tropicalis EST project 2001
 TROPICALIS_SEQUENCE_ID: TNeu072e23.plksp6
 Sequencing primer: SP6.
 ORIGIN
 location/Qualifiers
 1..594
 /organism="Xenopus tropicalis"
 /mol_type="mRNA"
 /db_xref="taxon:8364"
 /clone="TNeu072e23"
 /dev_stage="neurula"
 /lab_host="Escherichia coli DH10B"
 /clone_lib="XGC-neurula"
 /note="Vector: pCS107; Site_1: EcoRI; Site_2: NotI; cDNA
 was oligo dt primed from 5ug of poly A+ RNA from neurula.
 EcoRI-NotI cut cDNA was then ligated into pCS107 with
 EcoRI at the 5' end and NotI at the 3' end."
 ORIGIN

Query Match 84.8%; Score 17.8; DB 1; Length 594;
 Best Local Similarity 90.5%; Pred. No. 1.5e+03;
 Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 CAGGACACAGTGGCTTTGAC 21
 |||||
 Db 426 CAGGACACAGTGGCTTTGAC 406

RESULT 26
 CX421278/c 595 bp mRNA linear EST 24-AUG-2005
 LOCUS CX421278 XGI-XZG15551.fwd NIH_XGC-tropbas7 Xenopus tropicalis cDNA clone
 DEFINITION IMAGE:7531609 5', mRNA sequence.
 CX421278
 ACCESSION CX421278.2 GI:73751990
 VERSION EST.
 KEYWORDS Xenopus tropicalis (western clawed frog)
 SOURCE Xenopus tropicalis
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
 Xenopodinae; Xenopus; Silurana.
 REFERENCE 1 (bases 1 to 595)
 AUTHORS Richardson,P., Lucas,S., Rohbear,D., Dettler,J.C., Ng,D.C.,
 Brokstein,P. and Lindquist,E.A.
 TITLE DOE Joint Genome Institute Xenopus tropicalis EST project
 JOURNAL Unpublished (2004)
 COMMENT On Jan 6, 2005 this sequence version replaced gi:57201981.
 Other ESTs: XGI XZG15551.rev
 Contact: Lindquist,E.A., Richardson,P.
 DOE Joint Genome Institute
 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
 Tel: 925 296 5600
 Fax: 925 296 5710
 Email: cdna@jgi-psf.org
 Tissue Procurement: Richard M. Harland laboratory, University of
 California, Berkeley: http://tropicalis.berkeley.edu/home
 cDNA Library Preparation: Richard M. Harland laboratory, University
 of California, Berkeley
 DNA Sequencing: DOE Joint Genome Institute: http://www.jgi.doe.gov
 Clone Distribution: I.M.A.G.E. Consortium/LNL:
 http://image.llnl.gov
 Naming Conventions: EST name is generated by the concatenation of
 the XGI Clone Id and the direction of sequencing. The suffix '.fwd'
 indicates a forward sequencing read of the insert. It does not
 necessarily reflect the orientation of the insert.
 Plate: XZG 0161 row: n column: 23

High quality sequence stop: 528.

FEATURES

source

1..595
Location/Qualifiers
/organism="Xenopus tropicalis"
/mol_type="mRNA"
/db_xref="taxon:8364"
/clone="IMAGE:7531609"
/cissue_type="whole embryo"
/dev_stage="Gastrula (st. 10.5-12.5)"
/lab_host="E. coli XL1-Blue derivative, Stratagene Electropen-Blue"
/clone_lib="NH_XGC tropaeas7"
/note="vector: PCS108; Site 1: SalI; Site 2: NotI; Gastrula library constructed by Russell B. Fletcher in R. Harland's lab using poly A RNA and oligo dt primers (Invitrogen Superscript Plasmid System for cDNA Synthesis and Cloning). SalI (5' end) -NotI (3' end) cDNA was inserted into vector PCS108 (http://mcb.berkeley.edu/labs/harland/pages/plasmids.html)."

ORIGIN

Query Match 84.8%; Score 17.8; DB 9; Length 595;
Best Local Similarity 90.5%; Pred. No. 1.5e+03;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CAGGACACAGTGGCTTTGAC 21
|||
166 CAGGACACAGTGGCTTTGAC 146

RESULT 27

DR5G15T 626 bp DNA linear GSS 22-NOV-2002
LOCUS Dario rerio genomic clone DKX-5G15, genomic survey sequence.
ACCESSION AL733484
VERSION AL733484.1 GI:21350492
KEYWORDS GSS.
SOURCE Dario rerio (zebrafish)
ORGANISM Dario rerio

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.
1 (bases 1 to 626)
Humphrey,S.J., Huckle,E. and Hunt,S.E.
Submitted (06-JUN-2002) The Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail contact: humphrey@sanger.ac.uk Unpublished

COMMENT This sequence was generated from the T7 end of BAC 5G15. 5G15 is part of the Dariokey BAC Library created by R. Plastek and N.V. Keygene.
Further details: http://www.sanger.ac.uk/Projects/D_rerio/.

FEATURES

source

1..626
Location/Qualifiers
/organism="Dario rerio"
/mol_type="genomic DNA"
/db_xref="taxon:7955"
/clone="DKX-5G15"
/cissue_type="Testis"
/note="vector pindigobAC-536"

ORIGIN

Query Match 84.8%; Score 17.8; DB 14; Length 626;
Best Local Similarity 90.5%; Pred. No. 1.5e+03;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CAGGACACAGTGGCTTTGAC 21
|||
398 CAGGACACAGTGGCTTTGTC 418

RESULT 28

A0625517/c A0625517 643 bp DNA linear GSS 16-JUN-1999
LOCUS CITBI-EI-2653J24.TR CITBI-EI Homo sapiens genomic clone 2653J24,
DEFINITION genomic survey sequence.

ACCESSION A0625517
VERSION A0625517.1 GI:5087909
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 643)
Zhao,S., Adams,M.D., Nierman,W., Malek,J., Shiyaya,H., Simon,M. and Venter,J.C.
Use of BAC End Sequences from Caltech Libraries for Sequence-Ready Map Building
Unpublished (1997)

JOURNAL

COMMENT

Other GSSs: CITBI-EI-2653J24.TF
Contact: Shaying Zhao, William Nierman, Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850
Tel: 301 838 0200
Fax: 301 838 0208
Email: hbe@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC end search page:
http://www.tigr.org/tdb/hungen/bac_end_search/bac_end_search.html.
Seq primer: M13 Reverse
Class: BAC ends.

FEATURES

source

1..643
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="2653J24"
/sex="male"
/cell_type="sperm"
/clone_lib="CITBI-EI"
/note="Vector: pBelobAC11, Site_1: EcoRI, Site_2: EcoRI, Caltech Human BAC Library D"

ORIGIN

Query Match 84.8%; Score 17.8; DB 11; Length 643;
Best Local Similarity 90.5%; Pred. No. 1.5e+03;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CAGGACACAGTGGCTTTGAC 21
|||
494 CAGGACACAGTGGCTTATGAC 474

RESULT 29

AL642415 649 bp mRNA linear EST 07-NOV-2003
LOCUS AL642415 XGC-neurula Xenopus tropicalis cDNA clone TNeu025K14.5,
DEFINITION mRNA sequence.

ACCESSION AL642415
VERSION AL642415.2 GI:38221123
KEYWORDS EST.
SOURCE Xenopus tropicalis (western clawed frog)
ORGANISM Xenopus tropicalis

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae; Xenopodinae; Xenopus; Silurana.
1 (bases 1 to 649)
Croning,M.D.R., Ashurst,J.L., Taylor,R., Zorn,A.M. and Rogers,J.
Sanger Xenopus tropicalis EST project 2001 (11_2003)
Unpublished (2003)

COMMENT On Nov 7, 2001 this sequence version replaced gi:16794540.
Contact: Huckle E
Sanger Institute

Hinxton, Cambridgeshire, CB10 1SA, UK
 Email: trop@sanger.ac.uk
 This sequence is from a Xenopus Gene Collection (XGC) library
 constructed by Aaron M. Zorn.
 cDNA was oligo dt primed from 5ug of poly A+ RNA from neurula.
 EcoRI-NotI cut cDNA was then ligated into pCS107 with EcoRI at the
 5' end and NotI at the 3' end.
 Vector: pCS107; Site_1: EcoRI; Site_2: NotI
 Host: Escherichia coli DH10B
 Sanger Xenopus tropicalis EST project 2001
 TROPICALIS SEQUENCE ID: TNeu025K14.plkSP6
 Sequencing primer: SP6.
 Location/Qualifiers

FEATURES

source

1..649
 /organism="Xenopus tropicalis"
 /mol_type="mRNA"
 /db_xref="taxon:8364"
 /clone="TNeu025K14"
 /dev_stage="neurula"
 /lab_host="Escherichia coli DH10B"
 /clone_lib="XGC-neurula"
 /note="Vector: pCS107; Site_1: EcoRI; Site_2: NotI; cDNA
 was oligo dt primed from 5ug of poly A+ RNA from neurula.
 EcoRI-NotI cut cDNA was then ligated into pCS107 with
 EcoRI at the 5' end and NotI at the 3' end."

ORIGIN

Query Match 84.8%; Score 17.8; DB 1; Length 649;
 Best Local Similarity 90.5%; Pred. No. 1.5e+03;
 Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CAGGACACAGTGGCTTTTGAC 21
 |||||
 Db 420 CAGGACACAGTGGCTTTTGAC 400

RESULT 30
 CT346911/c 683 bp DNA linear GSS 03-NOV-2005
 LOCUS CT346911
 DEFINITION Sus scrofa genomic clone CH242-42B24, genomic survey sequence.
 ACCESSION CT346911
 VERSION CT346911.1 GI:79916515
 KEYWORDS GSS.
 SOURCE Sus scrofa (pig)
 ORGANISM Sus scrofa
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Suina; Suidae;
 Sus.

REFERENCE 1 (bases 1 to 683)
 HUMPHRAY, S.J., PLUMB, R.W. and DURHAM, J.L.
 TITLE Direct Submission
 JOURNAL Submitted (01-NOV-2005) The Sanger Institute, Wellcome Trust Genome
 Campus, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
 humphray@sanger.ac.uk Unpublished
 This sequence was generated from the T7 end of BAC 42B24. 42B24 is
 part of the CHORI-242 BAC Library created by P. de Jong. Further
 details: http://www.sanger.ac.uk/Projects/S_scrofa/.
 Location/Qualifiers

FEATURES

source

1..683
 /organism="Sus scrofa"
 /mol_type="genomic DNA"
 /db_xref="taxon:9623"
 /clone="CH242-42B24"
 /tissue_type="White blood cells"
 /note="Vector pTARBAC1.3_BamHI
 sex female"

ORIGIN

Query Match 84.8%; Score 17.8; DB 14; Length 683;
 Best Local Similarity 90.5%; Pred. No. 1.5e+03;
 Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CAGGACACAGTGGCTTTTGAC 21

|||||
 Db 228 CAGGACACAGAGGCTTTGTGC 208

RESULT 31
 AG144458/c 710 bp DNA linear GSS 08-JAN-2002
 LOCUS AG144458
 DEFINITION Pan troglodytes DNA, clone: RP43-005H09.TU, genomic survey
 sequence.

ACCESSION AG144458
 VERSION AG144458.1 GI:16674136
 KEYWORDS GSS.
 SOURCE Pan troglodytes (chimpanzee)
 ORGANISM Pan troglodytes
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homnidae; Pan.

REFERENCE 1
 FUJIYAMA, A., HATTORI, M., TOYODA, A., TAYLOR, T.D., YADA, T.,
 TOROKI, Y., WATANABE, H. and SAKAKI, Y.
 TITLE BAC end sequences of Library RP43-43
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 710)
 FUJIYAMA, A., HATTORI, M., TOYODA, A., TAYLOR, T.D., YADA, T.,
 TOROKI, Y., WATANABE, H. and SAKAKI, Y.
 TITLE Direct Submission

Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
 and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
 1-7-22 Suicho-chou, Tsukuba, Ibaraki, Japan
 (E-mail: chimbese@sc.riken.go.jp, URL: <http://hgp.gsc.riken.go.jp/>,
 Tel:81-45-503-9111, Fax:81-45-503-9170)
 Clones are derived from the chimpanzee BAC library RP43-43 This BAC
 end was generated during the R&D process and may have higher chance
 of clone tracking errors.
 PRIMERS

Sequencing: TU
 LIBRARY
 Vector : pBACe3.6
 R.Site 1 : EcoRI
 R.Site 2 : EcoRI.
 Location/Qualifiers
 1..710
 /organism="Pan troglodytes"
 /mol_type="genomic DNA"
 /db_xref="taxon:9598"
 /clone="RP43-005H09.TU"
 /sex="male"
 /cell_type="lymphocytes"
 /clone_lib="RP43-43 Chimpanzee Male BAC Library"

ORIGIN

Query Match 84.8%; Score 17.8; DB 14; Length 710;
 Best Local Similarity 90.5%; Pred. No. 1.5e+03;
 Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CAGGACACAGTGGCTTTTGAC 21
 |||||
 Db 539 CAGGACACAGTGGCTTTATGAC 519

RESULT 32
 AZ560569 712 bp DNA linear GSS 20-NOV-2000
 LOCUS RP43-23-206G21.TU RP43-23 Mus musculus genomic clone
 DEFINITION RP43-23-206G21, genomic survey sequence.
 ACCESSION AZ560569
 VERSION AZ560569.1 GI:11240389
 KEYWORDS GSS.

SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muridae; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 712)
 AUTHORS Zhao, S., Nierman, W., Feldblum, T., Malek, J., Shatman, S., Akniet, B., Levins, M., McGinn, S., Tesgaye, G., Geer, K., Krol, M., de Jong, P. and Fraser, C. M.
 TITLE Mouse BAC End Sequences from Library RPCI-23
 JOURNAL Unpublished (1999)
 COMMENT Other GSSs: RPCI-23-206G21.TV
 Contact: Shaying Zhao
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 0200
 Fax: 301 838 0208
 Email: szhaoc@tigr.org
 Clones are derived from the mouse BAC library RPCI-23. For BAC library availability, please contact Pieter de Jong (pieter@dejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (<http://bacpac.med.buffalo.edu/orderingframe.htm>) or from Resea ch Genetics (inforesgen.com). BAC end page: http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html
 Plate: 206 row: G column: 21
 Seq primer: SP6
 Class: BAC ends.

FEATURES
 source Location/Qualifiers
 1..712
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="RPCI-23-206G21"
 /sex="Female"
 /lab_host="DH10B"
 /clone_id="RPCI-23"
 /note="Organ: Kidney/Brain; Vector: pBACe3.6; Site: 1; EcoRI; Site 2: EcoRI; Female C57BL/6J mouse kidney and/or brain genomic DNA was isolated and partially digested with a combination of EcoRI and EcoRI Methylase. Size selected DNA was cloned into the pBACe3.6 vector at the EcoRI sites. The ligation products were transformed into DH10B electrocompetent cells (BRL Life Technologies)."

ORIGIN
 Query Match 84.8%; Score 17.8; DB 11; Length 712;
 Best Local Similarity 90.5%; Pred. No. 1.5e+03;
 Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CAGGACACAGTGGCTTTGAC 21
 |||||
 Db 645 CAGGACACAGTGGCTTTGAC 665

RESULT 33
 CX380870/c 788 bp mRNA linear EST 08-SEP-2005
 LOCUS JGI_XZT52439.fwd NIH_XGC_tropTad5 Xenopus tropicalis cDNA clone
 DEFINITION IMAGE:7627184 5', mRNA sequence.
 ACCESSION CX380870
 VERSION CX380870.2 GI:74291806
 KEYWORDS EST.
 SOURCE Xenopus tropicalis (western clawed frog)
 ORGANISM Xenopus tropicalis
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae; Xenopodinae; Xenopus; Silurana.
 1 (bases 1 to 788)
 Richardson, P., Lucas, S., Rokhsar, D., Dettler, J. C., Ng, D. C., Brokstein, P. and Lindquist, E. A.
 DOE Joint Genome Institute Xenopus tropicalis EST project Unpublished (2004)
 On Jan 5, 2005 this sequence version replaced gi:57149427.
 Other ESTs: JGI_XZT52439.rev
 Contact: Lindquist, E. A., Richardson, P.
 DOE Joint Genome Institute

2800 Mitchell Drive, Walnut Creek, CA 94598, USA
 Tel: 925 296 5600
 Fax: 925 296 5710
 Email: cdna@jgi-psf.org
 Tissue Procurement: Richard M. Harland Laboratory, University of California, Berkeley: <http://ctropicalis.berkeley.edu/home>
 cDNA Library Preparation: Richard M. Harland Laboratory, University of California, Berkeley
 DNA Sequencing: DOE Joint Genome Institute: <http://www.jgi.doe.gov>
 Clone Distribution: I.M.A.G.E. Consortium/LNL:
<http://image.lnl.gov>
 Naming Conventions: EST name is generated by the concatenation of the JGI Clone Id and the direction of sequencing. The suffix 'fwd' indicates a forward sequencing read of the insert. It does not necessarily reflect the orientation of the insert.
 Plate: XZT 0545 row: m column: 6
 High quality sequence stop: 786.

FEATURES
 source Location/Qualifiers
 1..788
 /organism="Xenopus tropicalis"
 /mol_type="mRNA"
 /db_xref="taxon:8364"
 /clone="IMAGE:7627184"
 /tissue_type="whole embryo"
 /dev_stage="Tadpole (st. 36-41)"
 /lab_host="E. coli XL1-Blue derivative, Stratagene Electrogen-Blue"
 /clone_id="NIH_XGC_tropTad5"
 /note="Vector: pCS108; Site 1: SalI; Site 2: NotI; Tadpole library constructed by Russell B. Fletcher in R. Harland's lab using poly A RNA and oligo dt primers (Invitrogen SuperScript Plasmid System for cDNA Synthesis and Cloning). SalI (5' end) -NotI (3' end) cDNA was inserted into vector pCS108
 (<http://mb.berkeley.edu/labs/harland/pages/plasmids.html>)"

ORIGIN
 Query Match 84.8%; Score 17.8; DB 9; Length 788;
 Best Local Similarity 90.5%; Pred. No. 1.5e+03;
 Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CAGGACACAGTGGCTTTGAC 21
 |||||
 Db 764 CAGGACACAGTGGCTTTGAC 744

RESULT 34
 CX443805/c 793 bp mRNA linear EST 07-JAN-2005
 LOCUS JGI_XZG9472.fwd NIH_XGC_tropCas7 Xenopus tropicalis cDNA clone
 DEFINITION IMAGE:7526634 5', mRNA sequence.
 ACCESSION CX443805
 VERSION CX443805.1 GI:57259637
 KEYWORDS EST.
 SOURCE Xenopus tropicalis (western clawed frog)
 ORGANISM Xenopus tropicalis
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae; Xenopodinae; Xenopus; Silurana.
 1 (bases 1 to 793)
 Richardson, P., Lucas, S., Rokhsar, D., Dettler, J. C., Ng, D. C., Brokstein, P. and Lindquist, E. A.
 DOE Joint Genome Institute Xenopus tropicalis EST project Unpublished (2004)
 Contact: Lindquist, E. A., Richardson, P.
 DOE Joint Genome Institute
 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
 Tel: 925 296 5600
 Fax: 925 296 5710
 Email: cdna@jgi-psf.org
 Tissue Procurement: Richard M. Harland Laboratory, University of California, Berkeley: <http://ctropicalis.berkeley.edu/home>

CDNA Library Preparation: Richard M. Harland Laboratory, University of California, Berkeley
 DNA Sequencing: DOE Joint Genome Institute: <http://www.jgi.doe.gov>
 Clone Distribution: I.M.A.G.E. Consortium/LNL:
<http://image.llnl.gov>
 Naming Conventions: EST name is generated by the concatenation of the JGI Clone id and the direction of sequencing. The suffix '.fwd' indicates a forward sequencing read of the insert. It does not necessarily reflect the orientation of the insert.
 Plate: XZG 0097 row: 0 column: 16
 High quality sequence stop: 756.
 Location/Qualifiers

FEATURES

source

```
1..793
/organism="Xenopus tropicalis"
/mol_type="mRNA"
/db_xref="taxon:8364"
/clone="IMAGE:7526634"
/tissue_type="whole embryo"
/dev_stage="Gastrula (st. 10.5-12.5)"
/lab_host="E. coli XL1-Blue derivative, Stratagene Electropen-Blue"
/clone_lib="NIH_XGC_tropcas7"
/notes="vector: PCS108; Site 1: SalI; Site 2: NotI; Gastrula library constructed by Russell B. Fletcher in R. Harland's lab using poly A RNA and oligo dt primers (Invitrogen Superscript Plasmid System for cDNA Synthesis and Cloning). SalI (5' end) -NotI (3' end) cDNA was inserted into vector PCS108 (http://mcb.berkeley.edu/labs/harland/pages/plasmids.html)"
```

ORIGIN

Query Match 84.8%; Score 17.8; DB 9; Length 793;
 Best Local Similarity 90.5%; Pred. No. 1.6e+03;
 Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 CAGGACACAGTGGCTTTGAC 21
 |||||
 Db 732 CAGGACACAGTGGCTTTGAC 712

RESULT 35
 CX444795/c 794 bp mRNA linear EST 07-JAN-2005
 LOCUS JGI_XZG10768.fwd NIH_XGC_tropcas7 Xenopus tropicalis cDNA clone
 DEFINITION IMAGE:7527773 5', mRNA sequence.
 ACCESSION CX444795
 VERSION CX444795.1 GI:57260627
 KEYWORDS EST.
 SOURCE Xenopus tropicalis (western clawed frog)
 ORGANISM Xenopus tropicalis
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae; Xenopodinae; Xenopus; Silurana.
 1 (bases 1 to 794)
 Richardson, P., Lucas, S., Rohrer, D., Dettler, J.C., Ng, D.C., Brockstein, P., and Lindquist, E.A.
 DOE Joint Genome Institute Xenopus tropicalis EST project
 Unpublished (2004)
 CONTACT: Lindquist, E.A., Richardson, P.
 DOE Joint Genome Institute
 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
 Tel: 925 296 5600
 Fax: 925 296 5710
 Email: cdna@jgi-psf.org
 Tissue Procurement: Richard M. Harland Laboratory, University of California, Berkeley: <http://tropicals.berkeley.edu/home>
 CDNA Library Preparation: Richard M. Harland Laboratory, University of California, Berkeley
 DNA Sequencing: DOE Joint Genome Institute: <http://www.jgi.doe.gov>
 Clone Distribution: I.M.A.G.E. Consortium/LNL:
<http://image.llnl.gov>
 Naming Conventions: EST name is generated by the concatenation of

the JGI Clone id and the direction of sequencing. The suffix '.fwd' indicates a forward sequencing read of the insert. It does not necessarily reflect the orientation of the insert.
 Plate: XZG 0113 row: 0 column: 3
 High quality sequence stop: 707.
 Location/Qualifiers

FEATURES

source

```
1..794
/organism="Xenopus tropicalis"
/mol_type="mRNA"
/db_xref="taxon:8364"
/clone="IMAGE:7527773"
/tissue_type="whole embryo"
/dev_stage="Gastrula (st. 10.5-12.5)"
/lab_host="E. coli XL1-Blue derivative, Stratagene Electropen-Blue"
/clone_lib="NIH_XGC_tropcas7"
/notes="vector: PCS108; Site 1: SalI; Site 2: NotI; Gastrula library constructed by Russell B. Fletcher in R. Harland's lab using poly A RNA and oligo dt primers (Invitrogen Superscript Plasmid System for cDNA Synthesis and Cloning). SalI (5' end) -NotI (3' end) cDNA was inserted into vector PCS108 (http://mcb.berkeley.edu/labs/harland/pages/plasmids.html)"
```

ORIGIN

Query Match 84.8%; Score 17.8; DB 9; Length 794;
 Best Local Similarity 90.5%; Pred. No. 1.6e+03;
 Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 CAGGACACAGTGGCTTTGAC 21
 |||||
 Db 776 CAGGACACAGTGGCTTTGAC 756

RESULT 36
 CX404457/c 796 bp mRNA linear EST 01-SRP-2005
 LOCUS JGI_XZT35930.fwd NIH_XGC_tropTads Xenopus tropicalis cDNA clone
 DEFINITION IMAGE:7611970 5', mRNA sequence.
 ACCESSION CX404457
 VERSION CX404457.2 GI:74091209
 KEYWORDS EST.
 SOURCE Xenopus tropicalis (western clawed frog)
 ORGANISM Xenopus tropicalis
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae; Xenopodinae; Xenopus; Silurana.
 1 (bases 1 to 796)
 Richardson, P., Lucas, S., Rohrer, D., Dettler, J.C., Ng, D.C., Brockstein, P., and Lindquist, E.A.
 DOE Joint Genome Institute Xenopus tropicalis EST project
 Unpublished (2004)
 On Jan 6, 2005 this sequence version replaced gi:57185153.
 Other ESTs: JGI_XZT35930.rev
 CONTACT: Lindquist, E.A., Richardson, P.
 DOE Joint Genome Institute
 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
 Tel: 925 296 5600
 Fax: 925 296 5710
 Email: cdna@jgi-psf.org
 Tissue Procurement: Richard M. Harland Laboratory, University of California, Berkeley: <http://tropicals.berkeley.edu/home>
 CDNA Library Preparation: Richard M. Harland Laboratory, University of California, Berkeley
 DNA Sequencing: DOE Joint Genome Institute: <http://www.jgi.doe.gov>
 Clone Distribution: I.M.A.G.E. Consortium/LNL:
<http://image.llnl.gov>
 Naming Conventions: EST name is generated by the concatenation of the JGI Clone id and the direction of sequencing. The suffix '.fwd' indicates a forward sequencing read of the insert. It does not necessarily reflect the orientation of the insert.
 Plate: XZT 0373 row: c column: 8

High quality sequence stop: 789.

Location/Qualifiers
1..796

/organism="Xenopus tropicalis"
/mol_type="mRNA"
/db_xref="taxon:8364"
/clone="IMAGE:7611970"
/tissue_type="whole embryo"
/dev_stage="Tadpole (st. 36-41)"
/lab_host="E. coli XL1-Blue derivative, Stratagene Electrogen-Blue"
/clone_idb="NIH_XGC_tropTad5"
/note="Vector: pCS108; Site 1: SalI; Site 2: NotI; Tadpole library constructed by Russell B. Fletcher in R. Harland's lab using poly A RNA and oligo dt primers (Invitrogen Superscript Plasmid System for cDNA Synthesis and Cloning). SalI (5' end) -NotI (3' end) cDNA was inserted into vector pCS108 (http://mcb.berkeley.edu/labs/harland/pages/plasmids.html)"

ORIGIN

Query Match 84.8%; Score 17.8; DB 9; Length 796;
Best Local Similarity 90.5%; Pred. No. 1.6e+03;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 CAGGACACAGTGGCTTTGAC 21
|||||
Db 429 CAGGACACAGTGGCTTTGAC 409

RESULT 37

CX439900/c

LOCUS JGI XZG8004 fwd NIH XGC_tropGas7 Xenopus tropicalis cDNA clone
DEFINITION IMAGE:7524924 5', mRNA sequence.
CX439900
CX439900.1 GI:57255732

ACCESSION

CX439900

VERSION

KEYWORDS

SOURCE

ORGANISM

Xenopus tropicalis (western clawed frog)

Xenopus tropicalis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
Xenopodinae; Xenopus; Silurana.
1 (bases 1 to 799)
Richardson, P., Lucas, S., Rokhsar, D., Dettler, J.C., Ng, D.C.,
Brokstein, P. and Lindquist, E.A.
DOE Joint Genome Institute Xenopus tropicalis EST project
Unpublished (2004)
Contact: Lindquist, E.A., Richardson, P.
DOE Joint Genome Institute
2800 Mitchell Drive, Walnut Creek, CA 94598, USA
Tel: 925 296 5600
Fax: 925 296 5710
Email: cdna@jgi-psf.org
Tissue Procurement: Richard M. Harland Laboratory, University of
California, Berkeley: http://tropicals.berkeley.edu/home
cDNA Library Preparation: Richard M. Harland Laboratory, University
of California, Berkeley
DNA Sequencing: DOE Joint Genome Institute: http://www.jgi.doe.gov
Clone Distribution: I.M.A.G.E. Consortium/LLNL:
http://image.llnl.gov
Nomenclature: EST name is generated by the concatenation of
the JGI clone id and the direction of sequencing. The suffix '.fwd'
indicates a forward sequencing read of the insert. It does not
necessarily reflect the orientation of the insert.
Plate: XZG 0081 row: h column: 10
High quality sequence stop: 791.
Location/Qualifiers
1..799

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

FEATURES

SOURCE

/organism="Xenopus tropicalis"
/mol_type="mRNA"
/db_xref="taxon:8364"

ORIGIN

Query Match 84.8%; Score 17.8; DB 9; Length 799;
Best Local Similarity 90.5%; Pred. No. 1.6e+03;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 CAGGACACAGTGGCTTTGAC 21
|||||
Db 778 CAGGACACAGTGGCTTTGAC 758

RESULT 38

CX407260/c

LOCUS JGI XZT62721 fwd NIH XGC_tropTad5 Xenopus tropicalis cDNA clone
DEFINITION IMAGE:7637291 5', mRNA sequence.
CX407260
CX407260.2 GI:74301053

ACCESSION

CX407260

VERSION

KEYWORDS

SOURCE

ORGANISM

Xenopus tropicalis (western clawed frog)

Xenopus tropicalis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
Xenopodinae; Xenopus; Silurana.
1 (bases 1 to 829)
Richardson, P., Lucas, S., Rokhsar, D., Dettler, J.C., Ng, D.C.,
Brokstein, P. and Lindquist, E.A.
DOE Joint Genome Institute Xenopus tropicalis EST project
Unpublished (2004)
On Jan 6, 2005 this sequence version replaced gi:57187960.
Other ESTs: JGI XZT62721.rev
Contact: Lindquist, E.A., Richardson, P.
DOE Joint Genome Institute
2800 Mitchell Drive, Walnut Creek, CA 94598, USA
Tel: 925 296 5600
Fax: 925 296 5710
Email: cdna@jgi-psf.org
Tissue Procurement: Richard M. Harland Laboratory, University of
California, Berkeley: http://tropicals.berkeley.edu/home
cDNA Library Preparation: Richard M. Harland Laboratory, University
of California, Berkeley
DNA Sequencing: DOE Joint Genome Institute: http://www.jgi.doe.gov
Clone Distribution: I.M.A.G.E. Consortium/LLNL:
http://image.llnl.gov
Nomenclature: EST name is generated by the concatenation of
the JGI clone id and the direction of sequencing. The suffix '.fwd'
indicates a forward sequencing read of the insert. It does not
necessarily reflect the orientation of the insert.
Plate: XZT 0653 row: b column: 9
High quality sequence stop: 828.
Location/Qualifiers
1..829

FEATURES

SOURCE

/organism="Xenopus tropicalis"
/mol_type="mRNA"
/db_xref="taxon:8364"
/clone="IMAGE:7637291"
/tissue_type="whole embryo"
/dev_stage="Tadpole (st. 36-41)"
/lab_host="E. coli XL1-Blue derivative, Stratagene Electrogen-Blue"

Electroten-Blue"
/clone_1lb="NIH_XGC_troptrads"
/note="Vector: pCS108; Site 1: SalI; Site 2: NotI; Tadpole
library constructed by Russell B. Fletcher in R. Harland's
lab using poly A RNA and oligo dt primers (Invitrogen
SuperScript Plasmid System for cDNA Synthesis and
cloning). SalI (5' end) -NotI (3' end) cDNA was inserted
into vector pCS108
(http://mcb.berkeley.edu/labs/harland/pages/plasmids.html)

ORIGIN

Query Match 84.8%; Score 17.8; DB 9; Length 829;
Best Local Similarity 90.5%; Pred. No. 1.6e+03;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 1 CAGGACACAGTGGCTTTGAC 21
|||||
381 CAGGACACAGTGGCTTTGAC 361

RESULT 39 850 bp mRNA linear EST 08-SEP-2005
LOCUS CX396192/c JGI XZT56263.fwd NIH XGC troptrads Xenopus tropicalis cDNA clone
DEFINITION IMAGE:7631020 5', mRNA sequence.

ACCESSION CX396192
VERSION CX396192.2 GI:74295269
KEYWORDS EST.
SOURCE Xenopus tropicalis (western clawed frog)
ORGANISM Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoes; Pipidae;
Xenopodinae; Xenopus; Silurana.
1 (bases 1 to 850)
Richardson, P., Lucas, S., Rokhsar, D., Dettler, J.C., Ng, D.C.,
Brookstein, P., and Lindquist, B.A.
DOE Joint Genome Institute Xenopus tropicalis EST project
Unpublished (2004)
On Jan 6, 2005 this sequence version replaced gi:57176876.
Other ESTs: JGI XZT56263.rev
Contact: Lindquist, B.A., Richardson, P.
DOE Joint Genome Institute
2800 Mitchell Drive, Walnut Creek, CA 94598, USA
Tel: 925 296 5600
Fax: 925 296 5710
Email: cdna@jgi-psf.org
Tissue Procurement: Richard M. Harland Laboratory, University of
California, Berkeley: http://tropicalis.berkeley.edu/home
cDNA library Preparation: Richard M. Harland Laboratory, University
of California, Berkeley
DNA Sequencing: DOE Joint Genome Institute: http://www.jgi.doe.gov
Clone Distribution: I.M.A.G.E. Consortium/LNL:
http://image.lnl.gov
Naming Conventions: EST name is generated by the concatenation of
the JGI Clone Id and the direction of sequencing. The suffix '.fwd'
indicates a forward sequencing read of the insert. It does not
necessarily reflect the orientation of the insert.
Plate: XZT 0585 row: m column: 2
High quality sequence stop: 829.
Location/Qualifiers

FEATURES
source

1..850
/organism="Xenopus tropicalis"
/mol_type="mRNA"
/db_xref="taxon:8364"
/clone="IMAGE:7631020"
/tissue_type="whole embryo"
/dev_stage="Tadpole (st. 36-41)"
/lab_host="E. coli XL1-Blue derivative, Stratagene
Electroten-Blue"
/clone_1lb="NIH_XGC_troptrads"
/note="Vector: pCS108; Site 1: SalI; Site 2: NotI; Tadpole
library constructed by Russell B. Fletcher in R. Harland's

lab using poly A RNA and oligo dt primers (Invitrogen
SuperScript Plasmid System for cDNA Synthesis and
cloning). SalI (5' end) -NotI (3' end) cDNA was inserted
into vector pCS108
(http://mcb.berkeley.edu/labs/harland/pages/plasmids.html)

ORIGIN

Query Match 84.8%; Score 17.8; DB 9; Length 850;
Best Local Similarity 90.5%; Pred. No. 1.6e+03;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 1 CAGGACACAGTGGCTTTGAC 21
|||||
129 CAGGACACAGTGGCTTTGAC 109

RESULT 40 874 bp DNA linear GSS 13-MAY-2003
LOCUS CC295265/c CH261-52A19_RML.1 CH261 Gallus gallus genomic clone CH261-52A19,
DEFINITION genomic survey sequence.
ACCESSION CC295265
VERSION CC295265.1 GI:3066706
KEYWORDS GSS.
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
1 (bases 1 to 874)
Kremetzki, C., Higgsindocham, J., Wylie, K., Carter, J., McPherson, J.,
Warren, W., Graves, T., Wardis, R., and Wilson, R.
Gallus gallus BAC End Reads
Unpublished (2003)
Contact: Richard K. Wilson
Genome Sequencing Center
Washington University School of Medicine
Email: submissions@wustl.wustl.edu
Insert Length: 18200 Std Error: 0.00
Seq primer: RML TACGACTCCTATGAGGAGA
Class: BAC ends
High quality sequence start: 33
High quality sequence stop: 589.
Location/Qualifiers

FEATURES
source

1..874
/organism="Gallus gallus"
/mol_type="genomic DNA"
/strain="Red Jungle Fowl"
/db_xref="taxon:9031"
/clone="CH261-52A19"
/sex="female"
/cell_line="UCDP001, indred 256"
/clone_1lb="CH261"
/note="Vector: pTARBAC2.1; Site 1: EcoRI; Site 2: EcoRI;
CH261 Female Chicken library - For library and clone
ordering information: http://www.chori.org/bacpac"

ORIGIN

Query Match 84.8%; Score 17.8; DB 12; Length 874;
Best Local Similarity 90.5%; Pred. No. 1.6e+03;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 1 CAGGACACAGTGGCTTTGAC 21
|||||
593 CTGACACAGTGGCTTTGTC 573

RESULT 41 883 bp mRNA linear EST 07-SEP-2005
LOCUS CX418837/c JGI XZ665774.fwd NIH XGC troptrads Xenopus tropicalis cDNA clone
DEFINITION IMAGE:757925 5', mRNA sequence.

ACCESSION CX418837
 VERSION CX418837.2 GI:74266898
 KEYWORDS EST.
 SOURCE Xenopus tropicalis (western clawed frog)
 ORGANISM Xenopus tropicalis
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
 Xenopodinae; Xenopus; Silurana.

REFERENCE 1 (bases 1 to 883)
 Richardson,P., Lucas,S., Rokhsar,D., Dettler,J.C., Ng,D.C.,
 Brokstein,P. and Lindquist,E.A.
 Title: DOB Joint Genome Institute Xenopus tropicalis EST project
 Journal: Unpublished (2004)
 Comment: On Jan 6, 2005 this sequence version replaced gi:57199540.
 Contact: Lindquist,E.A., Richardson,P.
 DOE Joint Genome Institute
 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
 Tel: 925 296 5600
 Fax: 925 296 5710
 Email: cdna@jgi-psf.org
 Tissue Procurement: Richard M. Harland laboratory, University of
 California, Berkeley: http://tropicals.berkeley.edu/home
 cDNA Library Preparation: Richard M. Harland Laboratory, University
 of California, Berkeley
 DNA Sequencing: DOE Joint Genome Institute: http://www.jgi.doe.gov
 Clone Distribution: I.M.A.G.E. Consortium/LLNL:
 http://image.llnl.gov
 Naming Conventions: EST name is generated by the concatenation of
 the JGI Clone Id and the direction of sequencing. The suffix '.fwd'
 indicates a forward sequencing read of the insert. It does not
 necessarily reflect the orientation of the insert.
 Plate: XZG 0685 row: 1 column: 3
 High quality sequence stop: 827.
 Location/Qualifiers
 1..883
 /organism="Xenopus tropicalis"
 /mol_type="mRNA"
 /db_xref="taxon:8364"
 /clone="IMAGE:7579925"
 /tissue_type="whole embryo"
 /dev_stage="Gastrula (st. 10.5-12.5)"
 /lab_host="E. coli XL1-Blue derivative, Stratagene
 Electrotren-Blue"
 /clone_1ib="NIH XGC tropGa87"
 /note="Vector: pCS108; Site 1: SalI; Site 2: NotI;
 Gastrula library constructed by Russell B. Fletcher in R.
 Harland's lab using poly A RNA and oligo dt primers
 (Invitrogen Superscript Plasmid System for cDNA Synthesis
 and Cloning). SalI (5' end) -NotI (3' end) cDNA was
 inserted into vector pCS108
 (http://mcb.berkeley.edu/labs/harland/pages/plasmids.html)"

ORIGIN
 Query Match 84.8%; Score 17.8; DB 9; Length 883;
 Best Local Similarity 90.5%; Pred. No. 1.6e+03;
 Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CAGGACACAGTGGCTTTGAC 21
 |||||
 Db 789 CAGGACACAGTGGCTTTGAC 769

RESULT 42
 LOCUS DR897380 895 bp mRNA linear EST 01-AUG-2005
 DEFINITION JGI_XZT42063.fwd NIH XGC_tropTads Xenopus tropicalis cDNA clone
 IMAGE:7617582 5', mRNA sequence.
 ACCESSION DR897380
 VERSION DR897380.1 GI:71586632
 KEYWORDS EST.
 SOURCE Xenopus tropicalis (western clawed frog)
 ORGANISM Xenopus tropicalis

REFERENCE 1 (bases 1 to 895)
 Richardson,P., Lucas,S., Rokhsar,D., Dettler,J.C., Ng,D.C.,
 Brokstein,P. and Lindquist,E.A.
 Title: DOB Joint Genome Institute Xenopus tropicalis EST project
 Journal: Unpublished (2004)
 Comment: Other ESTs: JGI XZT42063.rev
 Contact: Lindquist,E.A., Richardson,P.
 DOE Joint Genome Institute
 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
 Tel: 925 296 5600
 Fax: 925 296 5710
 Email: cdna@jgi-psf.org
 Tissue Procurement: Richard M. Harland laboratory, University of
 California, Berkeley: http://tropicals.berkeley.edu/home
 cDNA Library Preparation: Richard M. Harland Laboratory, University
 of California, Berkeley
 DNA Sequencing: DOE Joint Genome Institute: http://www.jgi.doe.gov
 Clone Distribution: I.M.A.G.E. Consortium/LLNL:
 http://image.llnl.gov
 Naming Conventions: EST name is generated by the concatenation of
 the JGI Clone Id and the direction of sequencing. The suffix '.fwd'
 indicates a forward sequencing read of the insert. It does not
 necessarily reflect the orientation of the insert.
 Plate: XZT 0437 row: m column: 4
 High quality sequence stop: 827.
 Location/Qualifiers
 1..895
 /organism="Xenopus tropicalis"
 /mol_type="mRNA"
 /db_xref="taxon:8364"
 /clone="IMAGE:7617582"
 /tissue_type="whole embryo"
 /dev_stage="Tadpole (st. 36-41)"
 /lab_host="E. coli XL1-Blue derivative, Stratagene
 Electrotren-Blue"
 /clone_1ib="NIH XGC tropTads"
 /note="Vector: pCS108; Site 1: SalI; Site 2: NotI; Tadpole
 library constructed by Russell B. Fletcher in R. Harland's
 lab using poly A RNA and oligo dt primers (Invitrogen
 Superscript Plasmid System for cDNA Synthesis and
 Cloning). SalI (5' end) -NotI (3' end) cDNA was inserted
 into vector pCS108
 (http://mcb.berkeley.edu/labs/harland/pages/plasmids.html)"

ORIGIN
 Query Match 84.8%; Score 17.8; DB 10; Length 895;
 Best Local Similarity 90.5%; Pred. No. 1.6e+03;
 Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CAGGACACAGTGGCTTTGAC 21
 |||||
 Db 461 CAGGACACAGTGGCTTTGAC 441

RESULT 43
 LOCUS DT398630 895 bp mRNA linear EST 25-AUG-2005
 DEFINITION JGI CABI233.fwd NIH XGC_tropOv1 Xenopus tropicalis cDNA clone
 IMAGE:7855953 5', mRNA sequence.
 ACCESSION DT398630
 VERSION DT398630.1 GI:73784827
 KEYWORDS EST.
 SOURCE Xenopus tropicalis (western clawed frog)
 ORGANISM Xenopus tropicalis
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
 Xenopodinae; Xenopus; Silurana.

REFERENCE 1 (bases 1 to 895)
 Richardson,P., Lucas,S., Rokhsar,D., Dettler,J.C., Ng,D.C.,

TITLE Broksrein, P. and Lindquist, E.A.
JOURNAL DOE Joint Genome Institute Xenopus tropicalis EST project
COMMENT Unpublished (2004)
 Other ESTs: JGI CAB12333.rev
 Contact: Lindquist, E.A., Richardson, P.
 DOE Joint Genome Institute
 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
 Tel: 925 296 5600
 Fax: 925 296 5710
 Email: cdna@jgi-psf.org
 Tissue Procurement: Robert M. Grainger
 CDNA Library Preparation: Bruce Blumberg Laboratory, University of California, Irvine
 CDNA Library Arrayed by: DOE Joint Genome Institute:
<http://www.jgi.doe.gov>
 DNA Sequencing: DOE Joint Genome Institute: <http://www.jgi.doe.gov>
 Clone Distribution: I.M.A.G.E. Consortium/LNL:
<http://image.llnl.gov>

Naming Conventions: EST name is generated by the concatenation of the JGI Clone Id and the direction of sequencing. The suffix '.fwd' indicates a forward sequencing read of the insert. It does not necessarily reflect the orientation of the insert.
 Plate: CAB1 0025 row: 1 column: 7
 High quality sequence stop: 842.
 Location/Qualifiers

FEATURES
 source

1. 895
 /organism="Xenopus tropicalis"
 /mol_type="mRNA"
 /strain="N6 (Nigerian 6th generation inbred)"
 /db_xref="taxon:8364"
 /clone="IMAGE:7855953"
 /sex="female"
 /tissue_type="Oviduct"
 /dev_stage="Adult"
 /lab_host="ElectroMAX DH10B T1 Phage Resistant cells"
 /clone_1fb="NIH XGC tropov11"
 /note="Vector: pCS107; Site 1: EcoRI; Site 2: XhoI; The library was prepared from 5 ug of poly A+ RNA by oligo-dT priming (5'-ACTAGTGGCGCCCTAGAGCTCGAGTTTCTTTTCTTTT-3') and Stratascript reverse transcriptase. After ligation of EcoRI adapters (5'-AATTCGACGACG-3') followed by kinasing adapters and by XhoI digestion, the cDNA was size selected by chromatography on Sepharose CL-2B columns and fractions containing cDNAs larger than 1000 bp were ligated into EcoRI/XhoI-digested pCS107. Reference for library construction: Current Genomics 4, 635-644. Library constructed by Michelle Tabb and Bruce Blumberg (Dept of Developmental and Cell Biology, University of California, Irvine)."

ORIGIN
 Query Match 84.8%; Score 17.8; DB 10; Length 895;
 Best Local Similarity 90.5%; Pred. No. 1.6e+03;
 Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CAGGACACAGTGGCTTTGAC 21
 |||||
 Db 476 CAGGACACAGTGGCTTTGAC 456

RESULT 44
 CX333579 896 bp mRNA linear EST 08-SEP-2005
 LOCUS CX333579 JGI XZT69314.fwd NIH XGC tropoXenopus tropicalis CDNA clone
 DEFINITION IMAGE:7788320 5', mRNA sequence.
 ACCESSION CX333579
 VERSION CX333579.2 GI:74307480
 KEYWORDS EST.
 SOURCE Xenopus tropicalis (western clawed frog)
 ORGANISM Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;

REFERENCE
 1 (bases 1 to 896)
 Richardson, P., Lucas, S., Rohrer, D., Deter, J.C., Ng, D.C., Broksrein, P., and Lindquist, E.A.
 DOE Joint Genome Institute Xenopus tropicalis EST project
 Unpublished (2004)
 On Jan 4, 2005 this sequence version replaced gi:57070051.
 Other ESTs: JGI XZT69314.rev
 Contact: Lindquist, E.A., Richardson, P.
 DOE Joint Genome Institute
 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
 Tel: 925 296 5600
 Fax: 925 296 5710
 Email: cdna@jgi-psf.org

Tissue Procurement: Richard M. Harland Laboratory, University of California, Berkeley: <http://tropicalis.berkeley.edu/home>
 CDNA Library Preparation: Richard M. Harland Laboratory, University of California, Berkeley
 DNA Sequencing: DOE Joint Genome Institute: <http://www.jgi.doe.gov>
 Clone Distribution: I.M.A.G.E. Consortium/LNL:
<http://image.llnl.gov>
 Naming Conventions: EST name is generated by the concatenation of the JGI Clone Id and the direction of sequencing. The suffix '.fwd' indicates a forward sequencing read of the insert. It does not necessarily reflect the orientation of the insert.
 Plate: XZT 0721 row: C column: 2
 High quality sequence stop: 803.
 Location/Qualifiers

FEATURES
 source

1. 896
 /organism="Xenopus tropicalis"
 /mol_type="mRNA"
 /db_xref="taxon:8364"
 /clone="IMAGE:7788220"
 /tissue_type="whole embryo"
 /dev_stage="Tadpole (st. 36-41)"
 /lab_host="B. coli XLI-Blue derivative, Stratagene Electrogen-Blue"
 /clone_1fb="NIH XGC tropoXds"
 /note="Vector: pCS108; Site 1: SalI; Site 2: NotI; Tadpole library constructed by Russell B. Fletcher in R. Harland's lab using poly A RNA and oligo dT primers (Invitrogen Superscript Plasmid System for cDNA Synthesis and cloning). SalI (5' end) -NotI (3' end) cDNA was inserted into vector pCS108
 (http://mcb.berkeley.edu/labs/harland/pages/plasmids.html)"

ORIGIN
 Query Match 84.8%; Score 17.8; DB 8; Length 896;
 Best Local Similarity 90.5%; Pred. No. 1.6e+03;
 Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CAGGACACAGTGGCTTTGAC 21
 |||||
 Db 785 CAGGACACAGTGGCTTTGAC 765

RESULT 45
 CF222676 906 bp mRNA linear EST 04-AUG-2003
 LOCUS CF222676 AGENCOURT 15066211 NICHD XGC EmbX Xenopus tropicalis CDNA clone
 DEFINITION IMAGE:6981253 5', mRNA sequence.
 ACCESSION CF222676
 VERSION CF222676.1 GI:33423384
 KEYWORDS EST.
 SOURCE Xenopus tropicalis (western clawed frog)
 ORGANISM Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae; Xenopodinae; Xenopus; Silurana.
 1 (bases 1 to 906)
 NIH-MGC <http://mhc.mcl.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL
COMMENT

Unpublished (1999)
Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NCI
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgsdb-remail.nih.gov
Tissue Procurement: Robert M. Grainger
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MCC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LLM14642 row: c column: 12
High quality sequence start: 4
High quality sequence stop: 715.
Location/Qualifiers

FEATURES
SOURCE

1..906
/organism="Xenopus tropicalis"
/mol_type="mRNA"
/db_xref="taxon:8364"
/clone="IMAGE:6981253"
/tissue_type="gastrula"
/dev_stage="embryo, stages 10-13"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NICHD XGC Emb5"
/note="Vector: pCMV-SF0R6.1; Site 1: NotI; Site 2: EcoRV;
Cloned unidirectionally. Primer: Oligo dT. Average insert
size 2.0 kb. Constructed by Invitrogen. Note: This is a
Xenopus Gene Collection (XGC) library."

ORIGIN

Query Match 84.8%; Score 17.8; DB 5; Length 906;
Best Local Similarity 90.5%; Pred. No. 1.6e+03;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CAGGACACAGTGGCTTTGAC 21
|||||
Db 755 CAGGACACAGTGGCTTTGAC 735

RESULT 46
CC304636 971 bp DNA linear GSS 13-MAY-2003
LOCUS CH261-89N15.RM1.1 CH261 Gallus gallus genomic clone CH261-89N15,
DEFINITION genomic survey sequence.
ACCESSION CC304636
VERSION CC304636.1 GI:30676077
KEYWORDS GSS.
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
1 (bases 1 to 971)
Kremetzki,C., Higgishotham,J., Wylie,K., Carter,J., McPherson,J.,
Warren,M., Graves,T., Mardis,E. and Wilson,R.
Gallus gallus BAC End Reads
Unpublished (2003)
Contact: Richard K. Wilson
Genome Sequencing Center
Washington University School of Medicine
Email: submissions@wustl.edu
Insert length: 182000 Std Error: 0.00
Seq primer: RM1 TACGACTGACTATAGGAGA
Class: BAC ends
High quality sequence start: 34
High quality sequence stop: 713.
Location/Qualifiers

FEATURES
SOURCE

1..971
/organism="Gallus gallus"
/mol_type="genomic DNA"
/strain="Red Jungle Fowl"

ORIGIN

Query Match 84.8%; Score 17.8; DB 12; Length 971;
Best Local Similarity 90.5%; Pred. No. 1.6e+03;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CAGGACACAGTGGCTTTGAC 21
|||||
Db 125 CAGGACACAGTGGCTTTGAC 145

RESULT 47
AK034662 2617 bp mRNA linear HTC 02-SEP-2005
LOCUS AK034662
DEFINITION Mus musculus 12 days embryo embryonic body between diaphragm region
and neck cDNA, RIKEN full-length enriched library, clone:1943002A14
product:hypothetical protein, full insert sequence.
ACCESSION AK034662
VERSION AK034662.1 GI:26330096
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Muridae; Murinae; Mus.

REFERENCE
AUTHORS Carninci,P. and Hayashizaki,Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
PUBMED 10349636
REFERENCE
AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
Itoh,M., Kono,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
PUBMED 11042159

REFERENCE
AUTHORS Shibata,K., Itoh,M., Aizawa,K., Nagao,S., Sasaki,N., Carninci,P.,
Kono,H., Akiyama,U., Nishi,K., Kitanaka,T., Tashiro,H., Itoh,M.,
Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishino,T., Harada,A.,
Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kasaiwa,K.,
Fujiwara,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Matsubara,M.,
Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura,S., Kawai,J.,
Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
TITLE RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
PUBMED 11076861

REFERENCE
AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the
PANTOM Consortium.
TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409, 685-690 (2001)
REFERENCE
AUTHORS The FANTOM Consortium, the RIKEN Genome Exploration Research Group
Phase I and II Team.
TITLE Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
JOURNAL Nature 420, 563-573 (2002)
REFERENCE
AUTHORS RIKEN Genome Exploration Research Group, Genome Science Group
(Genome Network Core Team) and the PANTOM Consortium.
TITLE Antisense Transcription in the Mammalian Transcriptome
JOURNAL Science 309, 1564-1566 (2005)

REFERENCE
AUTHORS
TITLE
JOURNAL
AUTHORS

7
The FANTOM Consortium, Riken Genome Exploration Research Group and
Genome Science Group (Genome Network Project Core Group).
The Transcriptional Landscape of the Mammalian Genome
Science 309, 1559-1563 (2005)
8 (bases 1 to 2617)

ADACHI,J., AIZAWA,K., AKIMURA,T., ARAKAWA,T., BONO,H., CARNINCI,P.,
FUKUDA,S., FURUNO,M., HANAGAKI,T., HARA,A., HASHIZUME,W.,
HAYASHIDA,K., HAYATEU,N., HIMOTO,K., HIRAKA,T., HITOZANE,T.,
HOTI,P., IMOTANI,K., ISHII,Y., ITOH,M., KAGAWA,I., KAKUKAWA,T.,
KATO,H., KAWAI,J., KOJIMA,Y., KONDO,S., KONNO,H., KOUNDA,M.,
KOYA,S., KURIHARA,C., MATSUYAMA,T., MIYAZAKI,A., MURATA,M.,
NAKAMURA,M., NISHI,K., NOMURA,K., NUMAZAKI,R., OHNO,M., OHSAITO,N.,
OKAZAKI,Y., SAITO,R., SAITOH,H., SAKAI,C., SAKAI,K., SAKAZUME,N.,
SANO,H., SAKAKI,D., SHIBATA,K., SHINGAWA,A., SHIRAKI,T.,
SOGABE,Y., TAGAMI,M., TAGAWA,A., TAKAHASHI,F., TAKAKU-AKASHI,S.,
TAKEDA,Y., TANAKA,T., TOMARU,A., TOYA,T., YASUNISHI,A.,
MURAMATSU,M. and HAYASHIZAKI,Y.

TITLE
JOURNAL

Submitted (16-JUN-2001) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa, 230-0045, Japan (E-mail:genome-res@gsc.riken.jp,
URL:http://genome.gsc.riken.jp/, Tel:81-45-503-9222,
Fax:81-45-503-9216)

COMMENT
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL:http://genome.gsc.riken.jp/
URL:http://fantom.gsc.riken.jp/
Location/Qualifiers

FEATURES
source

1..2617
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM_DB:9430022A14"
/db_xref="taxon:10090"
/clone="9430022A14"
/tissue_type="embryonic body between diaphragm region and
neck"
/clone_1ib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="12 days embryo"
146..1624
/note="unnamed protein product, hypothetical protein
(evidence: Procrest,decoder, longest-ORF)
putative"
/codon_start=1
/protein_id="BAC28787.1"
/db_xref="GI:26330097"
/translation="WRPPAPGAPRCLSMASLFCASASTLPGPMAAKWKEPILTL
KTASBARPGAGKGRKNSRTGATPEPPGKRNKSAELGSKSSGEPDPSRSD
SLQIIOAAKGRKNSRTVATAPTEPPGATVAGCTVYSILSLIVDQKNGKS
AYPGGCTSSSSSSSSSSSSSSSGPGLDRAKIRQDPTVILIEDYADPDAKRTG
QRDARLVGENDGYMEPYDAQMITEIRRGSDPLVKALQLDGSSEGEVYKATA
KRSSKDLGKRPOLYDTPPEBSGCGRAVEKTRPADSRLEBDEDRPAABYEDPEW
KREQIALSLVQPEGSDRPGEDAGRPDHQKTKPLSLDGHGDEKDPGLALEKOP
WYHGISRAEASRLQPCKEAYLVNRSSESGSKSYIALKTSOGCVHIIIVQTKONKY
TLNQTSNVFDSIPRYVHYYSNAKLPRKGAHEMTLHPVHKHA"

ORIGIN

Query Match 84.8%; Score 17.8; DB 6; Length 2617;
Beet Local Similarity 90.5%; Pired. No. 1.8e+03;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CAGGACACAGTGGCTTTGAC 21
|||||
DB 1756 CAGGACACAGTGGCTTTGAC 1776
|||||

RESULT 48
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AK157580
Mus musculus activated spleen cDNA, RIKEN full-length enriched
library, clone:FB030224M15 product:hypothetical protein, full insert
sequence.
AK157580.1 GI:74185930
HTC, CAP trapper.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridea; Muridae; Murinae; Mus.

REFERENCE
AUTHORS
TITLE
JOURNAL
PUBMED

Carninci,P. and Hayashizaki,Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
10349636

REFERENCE
AUTHORS
TITLE
JOURNAL
PUBMED

Carninci,P., Shibata,Y., Hayateu,N., Sugahara,Y., Shibata,K.,
Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
11042159

REFERENCE
AUTHORS
TITLE
JOURNAL
PUBMED

Shibata,K., Itoh,M., Aizawa,K., Nagaoaka,S., Sasaki,N., Carninci,P.,
Konno,H., Akiyama,J., Nishi,K., Kitsuana,T., Tashiro,T., Itoh,M.,
Suni,N., Ishii,Y., Nakamura,S., Hazama,M., Nishino,H., Harada,A.,
Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watanishi,M.,
Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura,S., Kawai,J.,
Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system-384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
11076861

REFERENCE
AUTHORS
TITLE
JOURNAL
PUBMED

Kawai,J., Shinagawa,A., Shibata,K., Yoshino,M., Itoh,M., Ishii,Y.,
Aizawa,T., Hara,A., Fukunishi,Y., Konno,H., Adachi,J., Fukuda,S.,
Aizawa,K., Izawa,M., Nishi,K., Kiyosawa,H., Kondo,S., Yamanaoka,I.,
Saito,T., Okazaki,Y., Gojobori,T., Bono,H., Kasukawa,T., Saito,R.,
Kadota,K., Matsuda,H., Ashburner,M., Batalov,S., Casavant,T.,
Pletschmann,W., Gaasterland,T., Gissi,C., King,B., Kochiwa,H.,
Kuehl,P., Lewis,S., Matsuo,Y., Nikaio,T., Pesole,G.,
Quackenbush,J., Schriml,L.M., Staubli,F., Suzuki,R., Tomita,M.,
Wagner,L., Washio,T., Sakai,K., Okido,T., Furuno,M., Aono,H.,
Baldarelli,R., Barsh,G., Blake,J., Boffelli,D., Bojunga,N.,
Carninci,P., de Bonaldo,M.F., Brownstein,M.J., Bult,C.,
Fletcher,C., Fujita,M., Gariboldi,M., Gustincich,S., Hill,D.,
Hofmann,M., Hume,D.A., Kamiya,M., Lee,N.H., Lyons,P.,
Marchionni,L., Mashima,J., Mazzaletti,J., Mombauts,P., Nordone,P.,
Rising,B., Ringwald,M., Rodriguez,I., Sakamoto,N., Sasaki,H.,
Sato,K., Schonbach,C., Seya,T., Shibata,Y., Storch,K.F., Suzuki,H.,
Toyo-Oka,K., Wang,K.H., Weitz,C., Whiteaker,C., Wilmink,L.,
Wyshaw-Boris,A., Yoshida,K., Haegawa,Y., Kawaji,H., Kortsuki,S.
and Hayashizaki,Y.

RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium
Functional annotation of a full-length mouse cDNA collection
Nature 409 (6821), 685-690 (2001)
11217851

REFERENCE
AUTHORS
TITLE
JOURNAL
PUBMED

Okazaki,Y., Furuno,M., Kasukawa,T., Adachi,J., Bono,H., Kondo,S.,
Nikaido,I., Ohsato,N., Saito,R., Suzuki,H., Yamanaoka,I.,
Kiyosawa,H., Yagi,K., Tomaru,Y., Haasgawa,Y., Nogami,A.,
Schonbach,C., Gojobori,T., Baldarelli,R., Hill,D.P., Bult,C.,
Hume,D.A., Quackenbush,J., Schriml,L.M., Kanapin,A., Matsuda,H.,
Batalov,S., Beisel,K.W., Blake,J.A., Bradt,D., Brusic,V.,
Chochia,C., Corradi,L.R., Cousins,S., Daille,B., Dragan,T.A.,
Fletcher,C.F., Forrest,A., Frazer,K.S., Gaasterland,T.,
Gariboldi,M., Gissi,C., Godzik,A., Gough,J., Grimmond,S.,

CONSRMT TITLE JOURNAL PUBMED REFERENCES AUTHORS	CONSRMT TITLE JOURNAL PUBMED REFERENCES AUTHORS	CONSRMT TITLE JOURNAL PUBMED REFERENCES AUTHORS	CONSRMT TITLE JOURNAL PUBMED REFERENCES AUTHORS
Guatlich,S., Hirokawa,N., Jackson,I.J., Jarvis,E.D., Kanai,A., Kutachin,I.V., Dee,Y., Lenhard,B., Lyons,P.A., Maglott,D.R., Maltala,L., Marchionni,L., McKenzie,L., Miki,H., Nagashima,T., Nunata,K., Okido,T., Pavan,M.J., Petter,G., Pesole,G., Petrovsky,N., Pillai,R., Pontius,J.U., Qi,D., Ramachandran,S., Ravasi,T., Reed,J.C., Reed,D.J., Reid,J., Ring,B.Z., Ringwald,M., Sandelin,A., Schneider,C., Sempke,C.A., Setou,M., Shimada,K., Sultana,R., Takanaka,Y., Taylor,M.S., Teasdale,R.D., Tomita,M., Verardo,R., Wagner,L., Walstedt,C., Wang,Y., Watanabe,Y., Wells,C., Wilming,L.G., Wyszewski-Boris,A., Yangisawa,M., Yang,I., Yang,L., Yuan,Z., Zavalan,M., Zhu,Y., Zimmer,A., Carrincci,P., Hayatenu,N., Hirozane-Kishikawa,T., Kono,H., Nakamura,M., Sakazume,N., Sato,K., Shiraki,T., Waki,K., Kawai,J., Aizawa,K., Arakawa,T., Fukuda,S., Hara,A., Hashizume,W., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Miyazaki,A., Sakai,K., Saeki,D., Shibata,K., Shinagawa,A., Yasunishi,A., Yoshino,M., Waterston,R., Lander,E.S., Rogers,J., Birney,E. and Hayashizaki,Y.	CONSRMT TITLE JOURNAL PUBMED REFERENCES AUTHORS	CONSRMT TITLE JOURNAL PUBMED REFERENCES AUTHORS	CONSRMT TITLE JOURNAL PUBMED REFERENCES AUTHORS
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs Nature 420 (6915), 563-573 (2002)	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs Nature 420 (6915), 563-573 (2002)	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs Nature 420 (6915), 563-573 (2002)	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs Nature 420 (6915), 563-573 (2002)
12466851	12466851	12466851	12466851
Carrincci,P., Kaenawa,T., Katayama,S., Gough,J., Frith,M.C., Mada,N., Oyama,R., Ravasi,T., Lenhard,B., Wells,C., Kodzima,R., Shimokawa,K., Bajic,V.B., Brenner,S.E., Batalov,S., Forrest,A.R., Zavalan,M., Davis,M.J., Wilming,L.G., Aldred,J., Allen,J.E., Amesl,Impionbarco,A., Apweiler,R., Alturaiya,R.N., Bailey,T.L., Barnes,M., Baxter,L., Beisel,K.W., Berano,T., Bono,H., Chak,A.M., Chiu,K.P., Choudhary,V., Christofidele,A., Clutterbuck,D.R., Crowe,M.L., Dalla,E., Dalrymple,B.P., De Bono,B., Della Gatta,G., di Bernardo,D., Down,T., Engstrom,P., Fagioli,M., Faulkner,G., Fletcher,C.F., Fukushima,T., Furuno,M., Futaki,S., Gariboldi,M., Georgii-Hemming,P., Gingeras,T.R., Gojobori,T., Green,R.E., Guatlich,S., Harbers,M., Hayashi,Y., Henach,T.K., Hirokawa,N., Hill,D., Hummel,L., Iacono,M., Iaco,K., Iwama,A., Ishikawa,T., Jact,M., Kanapin,A., Katoh,M., Kawasawa,Y., Kelso,J., Kitanura,H., Kitano,H., Kollas,G., Krishnan,S.P., Kruger,A., Kummerfeld,S.K., Kutachin,I.V., Lareau,L.F., Lazarevic,D., Lipovich,L., Liu,J., Liu,S., McWilliam,S., Madan,Babu,M., Madera,M., Marchionni,L., Matsuda,H., Matsuzawa,S., Miki,H., Mignone,F., Miyake,S., Morris,K., Mortu-Tabar,S., Mulder,N., Nakano,N., Nakachi,H., Ng,P., Nilsson,R., Nishiguchi,S., Nishikawa,S., Nori,F., Ohara,O., Okazaki,Y., Orlando,V., Pang,K.C., Pavan,M.J., Pavese,G., Pesole,G., Petrovsky,N., Piazza,S., Reed,J., Reid,J.F., Ring,B.Z., Ringwald,M., Roost,B., Ruan,Y., Salzberg,S.L., Sandelin,A., Schneider,C., Schonbach,C., Sekiguchi,K., Sempke,C.A., Seno,D., Sessa,L., Sheng,Y., Shibata,Y., Shimada,H., Shimada,K., Silva,D., Sincilar,B., Sperling,S., Stupka,E., Sugita,K., Sultana,R., Takanaka,Y., Taki,K., Tammoja,K., Tan,S.L., Tang,S., Taylor,M.S., Tegener,J., Teichmann,S.A., Ueda,H.R., van Nimwegen,E., Verardo,R., Wei,C.L., Yeai,K., Yamashita,H., Zabarovsky,B., Zhu,S., Zimmer,A., Hilde,M., Bult,C., Grimmond,S.M., Teasdale,R.D., Liu,E.T., Brusic,V., Quackenbush,J., Walstedt,C., Mattick,J.S., Hume,D.A., Kai,C., Saeki,D., Tomaru,Y., Fukuda,S., Kanamori-Katayama,M., Sunkai,M., Aoki,J., Arakawa,T., Iida,J., Imanura,K., Itoh,M., Kato,T., Kawai,J., Kawagashita,N., Kawashima,T., Kojima,M., Kondo,S., Kono,H., Nakano,K., Ninomiya,N., Nishio,T., Okada,M., Plessey,C., Shibata,K., Shiraki,T., Suzuki,S., Tagami,M., Waki,K., Watanishi,A., Okamura-Ohno,Y., Suzuki,H., Kawai,J., Hayashizaki,Y.	Carrincci,P., Kaenawa,T., Katayama,S., Gough,J., Frith,M.C., Mada,N., Oyama,R., Ravasi,T., Lenhard,B., Wells,C., Kodzima,R., Shimokawa,K., Bajic,V.B., Brenner,S.E., Batalov,S., Forrest,A.R., Zavalan,M., Davis,M.J., Wilming,L.G., Aldred,J., Allen,J.E., Amesl,Impionbarco,A., Apweiler,R., Alturaiya,R.N., Bailey,T.L., Barnes,M., Baxter,L., Beisel,K.W., Berano,T., Bono,H., Chak,A.M., Chiu,K.P., Choudhary,V., Christofidele,A., Clutterbuck,D.R., Crowe,M.L., Dalla,E., Dalrymple,B.P., De Bono,B., Della Gatta,G., di Bernardo,D., Down,T., Engstrom,P., Fagioli,M., Faulkner,G., Fletcher,C.F., Fukushima,T., Furuno,M., Futaki,S., Gariboldi,M., Georgii-Hemming,P., Gingeras,T.R., Gojobori,T., Green,R.E., Guatlich,S., Harbers,M., Hayashi,Y., Henach,T.K., Hirokawa,N., Hill,D., Hummel,L., Iacono,M., Iaco,K., Iwama,A., Ishikawa,T., Jact,M., Kanapin,A., Katoh,M., Kawasawa,Y., Kelso,J., Kitanura,H., Kitano,H., Kollas,G., Krishnan,S.P., Kruger,A., Kummerfeld,S.K., Kutachin,I.V., Lareau,L.F., Lazarevic,D., Lipovich,L., Liu,J., Liu,S., McWilliam,S., Madan,Babu,M., Madera,M., Marchionni,L., Matsuda,H., Matsuzawa,S., Miki,H., Mignone,F., Miyake,S., Morris,K., Mortu-Tabar,S., Mulder,N., Nakano,N., Nakachi,H., Ng,P., Nilsson,R., Nishiguchi,S., Nishikawa,S., Nori,F., Ohara,O., Okazaki,Y., Orlando,V., Pang,K.C., Pavan,M.J., Pavese,G., Pesole,G., Petrovsky,N., Piazza,S., Reed,J., Reid,J.F., Ring,B.Z., Ringwald,M., Roost,B., Ruan,Y., Salzberg,S.L., Sandelin,A., Schneider,C., Schonbach,C., Sekiguchi,K., Sempke,C.A., Seno,D., Sessa,L., Sheng,Y., Shibata,Y., Shimada,H., Shimada,K., Silva,D., Sincilar,B., Sperling,S., Stupka,E., Sugita,K., Sultana,R., Takanaka,Y., Taki,K., Tammoja,K., Tan,S.L., Tang,S., Taylor,M.S., Tegener,J., Teichmann,S.A., Ueda,H.R., van Nimwegen,E., Verardo,R., Wei,C.L., Yeai,K., Yamashita,H., Zabarovsky,B., Zhu,S., Zimmer,A., Hilde,M., Bult,C., Grimmond,S.M., Teasdale,R.D., Liu,E.T., Brusic,V., Quackenbush,J., Walstedt,C., Mattick,J.S., Hume,D.A., Kai,C., Saeki,D., Tomaru,Y., Fukuda,S., Kanamori-Katayama,M., Sunkai,M., Aoki,J., Arakawa,T., Iida,J., Imanura,K., Itoh,M., Kato,T., Kawai,J., Kawagashita,N., Kawashima,T., Kojima,M., Kondo,S., Kono,H., Nakano,K., Ninomiya,N., Nishio,T., Okada,M., Plessey,C., Shibata,K., Shiraki,T., Suzuki,S., Tagami,M., Waki,K., Watanishi,A., Okamura-Ohno,Y., Suzuki,H., Kawai,J., Hayashizaki,Y.	Carrincci,P., Kaenawa,T., Katayama,S., Gough,J., Frith,M.C., Mada,N., Oyama,R., Ravasi,T., Lenhard,B., Wells,C., Kodzima,R., Shimokawa,K., Bajic,V.B., Brenner,S.E., Batalov,S., Forrest,A.R., Zavalan,M., Davis,M.J., Wilming,L.G., Aldred,J., Allen,J.E., Amesl,Impionbarco,A., Apweiler,R., Alturaiya,R.N., Bailey,T.L., Barnes,M., Baxter,L., Beisel,K.W., Berano,T., Bono,H., Chak,A.M., Chiu,K.P., Choudhary,V., Christofidele,A., Clutterbuck,D.R., Crowe,M.L., Dalla,E., Dalrymple,B.P., De Bono,B., Della Gatta,G., di Bernardo,D., Down,T., Engstrom,P., Fagioli,M., Faulkner,G., Fletcher,C.F., Fukushima,T., Furuno,M., Futaki,S., Gariboldi,M., Georgii-Hemming,P., Gingeras,T.R., Gojobori,T., Green,R.E., Guatlich,S., Harbers,M., Hayashi,Y., Henach,T.K., Hirokawa,N., Hill,D., Hummel,L., Iacono,M., Iaco,K., Iwama,A., Ishikawa,T., Jact,M., Kanapin,A., Katoh,M., Kawasawa,Y., Kelso,J., Kitanura,H., Kitano,H., Kollas,G., Krishnan,S.P., Kruger,A., Kummerfeld,S.K., Kutachin,I.V., Lareau,L.F., Lazarevic,D., Lipovich,L., Liu,J., Liu,S., McWilliam,S., Madan,Babu,M., Madera,M., Marchionni,L., Matsuda,H., Matsuzawa,S., Miki,H., Mignone,F., Miyake,S., Morris,K., Mortu-Tabar,S., Mulder,N., Nakano,N., Nakachi,H., Ng,P., Nilsson,R., Nishiguchi,S., Nishikawa,S., Nori,F., Ohara,O., Okazaki,Y., Orlando,V., Pang,K.C., Pavan,M.J., Pavese,G., Pesole,G., Petrovsky,N., Piazza,S., Reed,J., Reid,J.F., Ring,B.Z., Ringwald,M., Roost,B., Ruan,Y., Salzberg,S.L., Sandelin,A., Schneider,C., Schonbach,C., Sekiguchi,K., Sempke,C.A., Seno,D., Sessa,L., Sheng,Y., Shibata,Y., Shimada,H., Shimada,K., Silva,D., Sincilar,B., Sperling,S., Stupka,E., Sugita,K., Sultana,R., Takanaka,Y., Taki,K., Tammoja,K., Tan,S.L., Tang,S., Taylor,M.S., Tegener,J., Teichmann,S.A., Ueda,H.R., van Nimwegen,E., Verardo,R., Wei,C.L., Yeai,K., Yamashita,H., Zabarovsky,B., Zhu,S., Zimmer,A., Hilde,M., Bult,C., Grimmond,S.M., Teasdale,R.D., Liu,E.T., Brusic,V., Quackenbush,J., Walstedt,C., Mattick,J.S., Hume,D.A., Kai,C., Saeki,D., Tomaru,Y., Fukuda,S., Kanamori-Katayama,M., Sunkai,M., Aoki,J., Arakawa,T., Iida,J., Imanura,K., Itoh,M., Kato,T., Kawai,J., Kawagashita,N., Kawashima,T., Kojima,M., Kondo,S., Kono,H., Nakano,K., Ninomiya,N., Nishio,T., Okada,M., Plessey,C., Shibata,K., Shiraki,T., Suzuki,S., Tagami,M., Waki,K., Watanishi,A., Okamura-Ohno,Y., Suzuki,H., Kawai,J., Hayashizaki,Y.	Carrincci,P., Kaenawa,T., Katayama,S., Gough,J., Frith,M.C., Mada,N., Oyama,R., Ravasi,T., Lenhard,B., Wells,C., Kodzima,R., Shimokawa,K., Bajic,V.B., Brenner,S.E., Batalov,S., Forrest,A.R., Zavalan,M., Davis,M.J., Wilming,L.G., Aldred,J., Allen,J.E., Amesl,Impionbarco,A., Apweiler,R., Alturaiya,R.N., Bailey,T.L., Barnes,M., Baxter,L., Beisel,K.W., Berano,T., Bono,H., Chak,A.M., Chiu,K.P., Choudhary,V., Christofidele,A

CONSTRM RIKEN Genome Exploration Research Group
 TITLE Antisense transcription in the mammalian transcriptome
 JOURNAL Science 309 (5740), 1564-1566 (2005)
 PUBMED 16141073
 REFERENCE 8 (bases 1 to 5736)
 AUTHORS Akakawa,T., Caminini,P., Fukuda,S., Hashizume,W., Hayashida,K.,
 Horii,F., Iida,Y., Imanura,K., Imotani,K., Itoh,M., Kanagawa,S.,
 Kawai,J., Kojima,M., Kono,H., Murata,M., Nakamura,M., Niimiya,N.,
 Nishiyori,H., Nomura,K., Ono,M., Sakazume,N., Sano,H., Saeki,D.,
 Shibata,K., Shiraki,T., Tagami,M., Tagami,Y., Waki,K., Watanishi,A.,
 Muramatsu,M. and Hayashizaki,Y.
 COMMENT Direct Submission
 Submitted (30-MAR-2004) Yoshihide Hayashizaki, The Institute of
 Physical and Chemical Research (RIKEN), Laboratory for Genome
 Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
 RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
 Kanagawa, 220-0045, Japan (E-mail:genome-resgsc.riken.jp,
 URL:http://genome.gsc.riken.jp/, Tel:81-45-503-9222,
 Fax:81-45-503-9216)
 CDNA library was prepared and sequenced in Mouse Genome
 Encyclopedia Project of Genome Exploration Research Group in Riken
 Genomic Sciences Center and Genome Science Laboratory in RIKEN.
 Division of Experimental Animal Research in Riken contributed to
 prepare mouse tissues.
 Tissues were provided by Dr. John Todd (Dept. of Medical Genetics
 Wellcome Trust Centre for Molecular Mechanisms in Disease Wellcome
 Trust/MRC building Addenbrookes Hospital Cambridge) whose
 assistance we gratefully acknowledge.
 Please visit our web site for further details.
 URL:http://genome.gsc.riken.jp/
 URL:http://fantom.gsc.riken.jp/
 Location/Qualifiers
 1. 5736
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="M0D"
 /db_xref="RANTOM:DB:F830224M15"
 /db_xref="taxon:10090"
 /clone="F830224M15"
 /tissue_type="activated spleen"
 /clone_id="RIKEN full-length enriched mouse cDNA library"
 135. 1613
 /note="unnamed protein product; hypothetical protein
 (evidence: CDS)
 putative"
 /cdon_start=1
 /protein_id="BAB34127.1"
 /db_xref="GI:74185931"
 /translation="MRRPAPGACPRCKNSGTGAAPTPKGKNGASABIGASAKSPGKSRISLRD
 SLOGIOLAAAGKRKRNRKRYTATTEBPBGAGATKACGCTYSRLIKVPTORNGS
 AITGGGATSSSSSSSSASPSFSLGEEDLKAKIMKODVYILBTDADYDKRTG
 QRARERGENGDYMEPYDAQMTTEIRRRSKDPLVVALQLDGSPSGSTVYEATA
 KRSSKDYDLGKPPQLYDTPPEBGGQRVAEVTIRPADSLPEEDRPAAEYQPMEM
 KRQIARALSVCFEGSDRPGRGDAGSPHMOETLKPTLSDHDEGKVDGALKEQNP
 WMHGSIRAAESRLQPCKEAAYLVNNSGNSGYSLALKTSGCCHVHIVAQTKDNKY

KEYWORDS GSS.
SOURCE Canis familiaris (dog)
ORGANISM Canis familiaris
Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Canidae;
Canis.

REFERENCE 1 (bases 1 to 390)
AUTHORS Kirkness, E.F., Batra, V., Halpern, A.L., Levy, S., Remington, K.,
Ruech, D.B., Delcher, A.L., Pop, M., Wang, W., Frazer, C.M. and
Venter, J.C.
The dog genome: survey sequencing and comparative analysis
Science 301 (5641), 1898-1903 (2003)
14512627

TITLE
JOURNAL
PUBMED
COMMENT
Contact: Kirkness EF
The Institute for Genomic Research
Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,
Rockville, MD 20850, USA
Tel: 301-838-0200
Fax: 301-838-0208
Email: ekirkness@tigr.org
Class: shocgun.

FEATURES
source
Location/Qualifiers
1..390
/organism="Canis familiaris"
/mol_type="genomic DNA"
/strain="Standard Poodle"
/db_xref="taxon:9615"
/clone_lib="Dog Library"
/note="Site 1: BstXI; Libraries were prepared from
peripheral blood"

ORIGIN
Query Match 82.9%; Score 17.4; DB 12; Length 390;
Best Local Similarity 94.7%; Pred. No. 2.2e+03;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 CAGGACACAGTGGCTTTG 19
Db 256 CAGGACACAGTGGCTTTG 238

RESULT 50
LOCUS AQ435575/c 430 bp DNA linear GSS 31-MAR-1999
DEFINITION HS_3147_A1_A10_SBE8 RPCI-11 Human Male BAC library Homo sapiens
genomic clone Plate=723 Col=19 Row=A, genomic survey sequence.
ACCESSION AQ435575
VERSION AQ435575.1 GI:4546914
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
1 (bases 1 to 430)
Mahairas, G.G., Wallace, J.C., Smith, K., Swartzell, S., Holzman, T.,
Keller, A., Shaker, R., Furlong, J., Young, J., Zhao, S., Adams, M.D. and
Hood, L.
Sequence-tagged connectors: A sequence approach to mapping and
scanning the human genome
Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
10449764
Contact: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Clones are derived from the human BAC library RPCI-11. For BAC
library availability, please contact Pieter de Jong
(pieterdejong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm)

or from Resear h Genetics (info@resgen.com). BAC end Web Server:
<http://www.hrcg.washington.edu>
Plate: 723 Row: A Column: 19
Seq primer: SP6
Class: BAC ends
High quality sequence stop: 430.
Location/Qualifiers
1..430
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone_lib="Plate=723 Col=19 Row=A"
/sex="male"
/note="Vector: pBAC3.6; Site 1: EcoRI; Site 2: EcoRI;
Male blood DNA was isolated from one randomly chosen donor
and partially digested with a combination of EcoRI and
EcoRI methylase. Size selected DNA was cloned into the
pBAC3.6 vector at EcoRI sites"

FEATURES
source
Location/Qualifiers
1..430
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone_lib="Plate=723 Col=19 Row=A"
/sex="male"
/note="Vector: pBAC3.6; Site 1: EcoRI; Site 2: EcoRI;
Male blood DNA was isolated from one randomly chosen donor
and partially digested with a combination of EcoRI and
EcoRI methylase. Size selected DNA was cloned into the
pBAC3.6 vector at EcoRI sites"

ORIGIN
Query Match 82.9%; Score 17.4; DB 11; Length 430;
Best Local Similarity 94.7%; Pred. No. 2.2e+03;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2 AGGACACAGTGGCTTTGA 20
Db 96 AGGACACAGTGGCTTTGA 78

Search completed: May 24, 2006, 07:55:58
Job time : 1512.83 secs

THIS PAGE IS BLANK

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 24, 2006, 05:34:58 ; Search time 300.037 Seconds
(without alignments)
650.664 Million cell updates/sec

Title: US-10-601-913-5

Perfect score: 28

Sequence: 1 GAACAGCAATACCAACACGTTGTGTC 28

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 5244920 seqs, 3486124231 residues

Total number of hits satisfying chosen parameters: 10489840

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 1000 summaries

Database :

N_Geneseq.8:*

1: geneseqn1980s:*

2: geneseqn1980s:*

3: geneseqn2000s:*

4: geneseqn2001as:*

5: geneseqn2001bs:*

6: geneseqn2002as:*

7: geneseqn2002bs:*

8: geneseqn2003as:*

9: geneseqn2003bs:*

10: geneseqn2003cs:*

11: geneseqn2003ds:*

12: geneseqn2004as:*

13: geneseqn2004bs:*

14: geneseqn2005s:*

15: geneseqn2006s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match	Length	DB	ID	Description
1	28	100.0	28	2	AAT03309		Aat03309 HPV E6 ge
2	28	100.0	28	2	AAT72300		Aat72300 Human pap
3	28	100.0	28	2	AAT72335		Aat72335 Human pap
4	28	100.0	28	2	AAT72299		Aat72299 Human pap
5	28	100.0	28	2	AAT72334		Aat72334 Human pap
6	28	100.0	452	14	ADK26909		Adk26909 Human pap
7	28	100.0	456	2	ABO76223		Abg76223 Human pap
8	28	100.0	459	10	ADC37149		Adc37149 459bp HPV
9	28	100.0	477	10	ADF09607		Adf09607 Human pap
10	28	100.0	477	10	AAL54430		Aal54430 B6 gRNA
11	28	100.0	477	10	AAL54445		Aal54445 HPV16 E6
12	28	100.0	477	14	ADK26908		Adk26908 Human pap
13	28	100.0	477	14	AEK51113		Aek51113 Human pap
14	28	100.0	519	2	AAT31834		Aat31834 Human pap
15	28	100.0	543	10	ADF31984		Adf31984 Human pap
16	28	100.0	570	2	AAQ75470		Aaq75470 HPV16 E6/
17	28	100.0	712	1	AAK91600		Aak91600 Partial n
18	28	100.0	747	12	AD044067		Ado44067 Nucleotid

19	28	100.0	747	12	AD044065		Ado44065 Nucleotid
20	28	100.0	747	12	AD044061		Ado44061 Nucleotid
21	28	100.0	747	12	AD044063		Ado44063 Nucleotid
22	28	100.0	747	12	AD044069		Ado44069 Nucleotid
23	28	100.0	747	12	AD044071		Ado44071 Nucleotid
24	28	100.0	768	13	ADR47004		Adr47004 Human pap
25	28	100.0	768	15	AEF40156		Aef40156 Human pap
26	28	100.0	776	2	AAT14663		Aat14663 E6/E7 reg
27	28	100.0	779	2	AAK89756		Aak89756 Probe seq
28	28	100.0	790	2	AAQ29389		Aaq29389 DNA encod
29	28	100.0	801	2	AAT31833		Aat31833 Human pap
30	28	100.0	822	2	AAK78792		Aak78792 HPV fusio
31	28	100.0	822	2	AAK29781		Aak29781 Prot.DI/3
32	28	100.0	822	14	AED52633		Aed52633 Fusion pr
33	28	100.0	879	2	AAK78795		Aak78795 HPV fusio
34	28	100.0	879	2	AAK29784		Aak29784 CLYTA-B6
35	28	100.0	879	14	AED52641		Aed52641 Fusion pr
36	28	100.0	921	3	AAK09477		Aak09477 Human pap
37	28	100.0	939	6	AAD35101		Aad35101 Human pap
38	28	100.0	1000	2	AAK08627		Aak08627 HPV-16 fr
39	28	100.0	1005	1	AAK91784		Aak91784 DNA probe
40	28	100.0	1116	2	AAK78793		Aak78793 HPV fusio
41	28	100.0	1116	2	AAK29782		Aak29782 Prot.DI/3
42	28	100.0	1116	14	AED52637		Aed52637 Fusion pr
43	28	100.0	1173	2	AAK78797		Aak78797 HPV fusio
44	28	100.0	1173	2	AAK29786		Aak29786 CLYTA-B6R
45	28	100.0	1173	14	AED52645		Aed52645 Fusion pr
46	28	100.0	7840	5	AAK55127		Aak55127 Nucleotid
47	28	100.0	7902	2	AAT09847		Aat09847 Human pap
48	28	100.0	7904	2	AAT94724		Aat94724 Human pap
49	28	100.0	7904	2	AAT94723		Aat94723 Human pap
50	28	100.0	7904	2	AAK33881		Aak33881 HPV-16 ge
51	28	100.0	7904	2	AAK99946		Aak99946 Nucleotid
52	28	100.0	7904	2	AAK99952		Aak99952 Nucleotid
53	28	100.0	7904	8	ADA03127		Ada03127 HPV E6 se
54	28	100.0	7904	11	ADO58335		Ado58335 Human pap
55	28	100.0	7904	14	ADZ45647		Adz45647 Human pap
56	26.4	94.3	80	6	ABQ73529		Abq73529 HPV-PTM r
57	26.4	94.3	115	6	ABQ73534		Abq73534 HPV-PTM r
58	26.4	94.3	138	6	ABQ73533		Abq73533 HPV-PTM r
59	26.4	94.3	149	6	ABQ73530		Abq73530 HPV-PTM r
60	26.4	94.3	193	6	ABQ73532		Abq73532 HPV-PTM r
61	26.4	94.3	193	6	ABQ73531		Abq73531 HPV-PTM r
62	22	78.6	22	10	ADD36695		Add36695 Human pap
63	22	78.6	22	10	ADD36890		Add36890 Human pap
64	22	78.6	22	10	ADD21963		Add21963 HPV E6 ge
65	22	78.6	22	10	ADD22270		Add22270 HPV E6 ge
66	22	78.6	30	2	AAT67677		Aat67677 wild type
67	22	78.6	53	10	ADD36582		Add36582 Human pap
68	22	78.6	53	10	ADD22101		Add22101 HPV E6 ge
69	21	75.0	30	2	AAQ44289		Aaq44289 Sequence
70	21	75.0	30	2	AAQ44288		Aaq44288 Sequence
71	21	75.0	30	14	ADZ64770		Adz64770 HPV 16 E6
72	21	75.0	33	2	AAQ76094		Aaq76094 Human pap
73	20.4	72.9	30	2	AAT66768		Aat66768 Mutant E6
74	20	71.4	20	10	ADD36697		Add36697 Human pap
75	20	71.4	20	10	ADD36892		Add36892 Human pap
76	20	71.4	20	10	ADD22272		Add22272 HPV E6 ge
77	20	71.4	20	10	ADD21965		Add21965 HPV E6 ge
78	20	71.4	23	3	AAK09870		Aak09870 Human pap
79	20	71.4	23	12	ADL71271		Adl71271 Probe #6
80	20	71.4	23	13	ADK64332		Adk64332 Human pap
81	20	71.4	24	2	AAQ23125		Aaq23125 HPV16 pro
82	20	71.4	24	2	AAQ23126		Aaq23126 HPV16 pro
83	20	71.4	24	2	AAT29920		Aat29920 Human pap
84	20	71.4	24	2	AAT29921		Aat29921 Human pap
85	20	71.4	51	10	ADD36584		Add36584 Human pap
86	20	71.4	51	10	ADD22103		Add22103 HPV E6 ge
87	19.6	70.0	30000	9	ADA03008		Ada03008 Human CAL
88	19.6	70.0	30000	10	ADB72746		Adb72746 Human CAL
89	19.6	70.0	30000	10	ADC85488		Adc85488 Human CAL
90	19.6	70.0	30000	12	ADM74603		Adm74603 Human car
91	19	67.9	19	12	ADP99269		Adp99269 HPV16 E6

C	92	19	67.9	36	12	ADG00247	AdG00247 Nicotiana	165	18.4	65.7	542	6	AA153420	AA153420 Human pap
C	93	19	67.9	47	12	ADP99270	AdP99270 HPV16 B6	C 166	18.4	65.7	653	6	ABN61338	ABN61338 Human can
C	94	19	67.9	47	12	ADP99270	AdP99270 HPV16 B6	C 167	18.4	65.7	12879	6	ABN62230	ABN62230 Prostate
C	95	19	67.9	450	12	AD044104	Ad044104 Nucleotid	C 168	18.4	65.7	12880	11	ADN39609	ADN39609 Cancer/an
C	96	19	67.9	453	4	AA115124	AA115124 Probe #50	C 169	18.4	65.7	13202	4	AAK51828	AAK51828 Human pol
C	97	19	67.9	453	4	ABAS6878	AbAS6878 Human foe	C 170	18.4	65.7	58000	14	AD242284	AD242284 Human K10
C	98	19	67.9	453	4	AA136460	AA136460 Probe #51	C 171	18.2	65.0	131	4	AA126845	AA126845 Probe #16
C	99	19	67.9	453	4	ABAA6321	AbAA6321 Human bre	C 172	18.2	65.0	131	4	ABA75098	ABA75098 Human foe
C	100	19	67.9	453	4	AAK04882	AAK04882 Human bre	C 173	18.2	65.0	131	4	AA155634	AA155634 Probe #24
C	101	19	67.9	453	5	AA104876	AA104876 Probe #48	C 174	18.2	65.0	131	4	AAK49738	AAK49738 Human bon
C	102	19	67.9	533	4	AAK42809	AAK42809 Human G P	C 175	18.2	65.0	131	4	AAK23614	AAK23614 Human bra
C	103	19	67.9	534	4	ABAA69015	AbAA69015 Human foe	C 176	18.2	65.0	131	4	ABSA49367	ABSA49367 Human liv
C	104	19	67.9	534	4	AA149205	AA149205 Probe #17	C 177	18.2	65.0	131	6	ABSA32330	ABSA32330 Human gen
C	105	19	67.9	534	4	ABAA51025	AbAA51025 Human bre	C 178	18.2	65.0	417	6	ABN95008	ABN95008 Gene #150
C	106	19	67.9	534	4	AAK43128	AAK43128 Human bon	C 179	18.2	65.0	466	10	ADBS81762	ADBS81762 Arabidops
C	107	19	67.9	534	4	ABSA42760	ABSA42760 Human liv	C 180	18.2	65.0	514	4	AA117640	AA117640 Probe #75
C	108	19	67.9	534	5	AA109498	AA109498 Probe #94	C 181	18.2	65.0	514	4	ABAA62577	ABAA62577 Human foe
C	109	19	67.9	534	4	AA124453	AA124453 Probe #14	C 182	18.2	65.0	514	4	AA142565	AA142565 Probe #11
C	110	19	67.9	924	4	ABAA51552	AbAA51552 Human bre	C 183	18.2	65.0	514	4	AAK36788	AAK36788 Human bon
C	111	19	67.9	924	4	ABAA36526	AbAA36526 Probe #14	C 184	18.2	65.0	514	4	AAK10934	AAK10934 Human bra
C	112	19	67.9	924	6	ABAA17863	AbAA17863 Human gen	C 185	18.2	65.0	514	4	ABSA36451	ABSA36451 Human liv
C	113	19	67.9	1023	6	AAQ48575	AAQ48575 HPV B6/77	C 186	18.2	65.0	514	6	ABSA10792	ABSA10792 Human gen
C	114	19	67.9	1069	12	AD080744	Ad080744 Porcine e	C 187	18.2	65.0	531	6	ABV95791	ABV95791 Human pan
C	115	19	67.9	1961	4	AA115278	AA115278 Probe #52	C 188	18.2	65.0	534	6	ABV95677	ABV95677 Human pan
C	116	19	67.9	1961	4	ABAA6452	AbAA6452 Human bre	C 189	18.2	65.0	534	6	ABV95677	ABV95677 Human pan
C	117	19	67.9	1961	4	ABAA26658	AbAA26658 Probe #51	C 190	18.2	65.0	632	3	AAK38756	AAK38756 Arabidops
C	118	19	67.9	1961	6	ABSA05364	ABSA05364 Human gen	C 191	18.2	65.0	1239	12	AD155282	AD155282 Human big
C	119	19	67.9	2353	4	AA119563	AA119563 Probe #94	C 192	18.2	65.0	1239	12	AD155282	AD155282 Human pol
C	120	19	67.9	2353	4	ABAA64582	AbAA64582 Human foe	C 193	18.2	65.0	1276	2	AAV79126	AAV79126 HPTT-1 co
C	121	19	67.9	2353	4	AA144754	AA144754 Probe #13	C 194	18.2	65.0	1695	3	AAK34705	AAK34705 Arabidops
C	122	19	67.9	2353	4	ABAA6708	AbAA6708 Human bre	C 195	18.2	65.0	2200	6	AB163311	AB163311 Breast ca
C	123	19	67.9	2353	4	ABAA31710	AbAA31710 Probe #10	C 196	18.2	65.0	2200	6	AB162890	AB162890 Breast ca
C	124	19	67.9	2353	4	AAK38756	AAK38756 Human bon	C 197	18.2	65.0	2200	6	ABK84351	ABK84351 Human CDN
C	125	19	67.9	2353	4	AAK33028	AAK33028 Human bra	C 198	18.2	65.0	2200	6	ABN96871	ABN96871 Gene #336
C	126	19	67.9	2353	4	ABSA38334	ABSA38334 Human liv	C 199	18.2	65.0	2441	5	AAK98926	AAK98926 Human pan
C	127	19	67.9	2353	5	AA105284	AA105284 Probe #52	C 200	18.2	65.0	2853	8	ACA37230	ACA37230 Prokaryot
C	128	19	67.9	2353	8	ABSA12830	ABSA12830 Human gen	C 201	18.2	65.0	2916	10	ADDB9788	ADDB9788 Human prp
C	129	19	67.9	4230	10	ADBB69573	ADBB69573 C. neofo	C 202	18.2	65.0	3357	15	ABE86404	ABE86404 Hepatic s
C	130	19	67.9	4230	8	ABX12953	ABX12953 DNA encod	C 203	18.2	65.0	3357	10	ADG90875	ADG90875 Human hyp
C	131	19	67.9	4230	14	ADAA06265	ADAA06265 DNA encod	C 204	18.2	65.0	4342	8	AAZ51554	AAZ51554 Human gen
C	132	19	67.9	4352	14	ABED22627	ABED22627 Human mut	C 205	18.2	65.0	4342	8	ABSA6338	ABSA6338 Human gen
C	133	19	67.9	4895	10	ADG62900	ADG62900 Human MCH	C 206	18.2	65.0	4394	12	ADN03661	ADN03661 Antipso
C	134	19	67.9	4895	12	ADQ24520	ADQ24520 Human PRO	C 207	18.2	65.0	4394	12	ADP20794	ADP20794 Human pro
C	135	19	67.9	4895	12	ADPE6691	ADPE6691 Human mls	C 208	18.2	65.0	4394	14	ADX08167	ADX08167 Cyclin-de
C	136	19	67.9	4895	14	ADY53442	ADY53442 Human MTH	C 209	18.2	65.0	4626	8	ACCA6213	ACCA6213 Human dit
C	137	19	67.9	4895	14	ABEA23877	ABEA23877 Human PRO	C 210	18.2	65.0	5099	14	ADZ80561	ADZ80561 Protein t
C	138	19	67.9	4895	14	ABED22625	ABED22625 Human mut	C 211	18.2	65.0	5289	5	ADL63192	ADL63192 Human ova
C	139	19	67.9	4895	14	ABED22626	ABED22626 Human mut	C 212	18.2	65.0	110000	14	ABEA39175_04	ABEA39175_04 Cont
C	140	19	67.9	4999	10	ADBB69212	ADBB69212 C. neofo	C 213	18.2	65.0	243335	14	ABEA42401_04	ABEA42401_04 Cont
C	141	19	67.9	5130	11	ACNA4919	ACNA4919 Human mRN	C 214	18.2	65.0	256294	13	ABD33020	ABD33020 Mouse can
C	142	19	67.9	5210	11	ACNA4919	ACNA4919 Breast ca	C 215	18.2	65.0	295644	14	ABEA35721	ABEA35721 L. pneumo
C	143	19	67.9	5212	3	AACT6356	AACT6356 Human ORF	C 216	18.2	65.0	295644	14	ABEA35721	ABEA35721 L. pneumo
C	144	19	67.9	8307	6	ABK90804	ABK90804 DNA encod	C 217	18.2	65.0	298667	12	AAAT40319	AAAT40319 L. prim
C	145	19	67.9	32134	4	AA103615	AA103615 Human rep	C 218	18.2	64.3	18	2	AAK33884	AAK33884 PCR prime
C	146	19	67.9	32134	4	ABAA07813	ABAA07813 Human ova	C 219	18.2	64.3	289	2	AAK40091	AAK40091 Gastric c
C	147	19	67.9	32134	4	AA103616	AA103616 Human rep	C 220	18.2	64.3	918	8	ACA29376	ACA29376 Prokaryot
C	148	19	67.9	32191	4	ABAA07814	ABAA07814 Human ova	C 221	18.2	64.3	918	8	ACA29376	ACA29376 Prokaryot
C	149	19	67.9	55114	14	ACNA4918	ACNA4918 Human gen	C 222	18.2	64.3	3001	14	ADZ61200	ADZ61200 Murine Th
C	150	19	67.9	110000	11	ADZ13651_1	ADZ13651_1 Cont	C 223	18.2	64.3	21921	4	AB105352	AB105352 Drosophi
C	151	19	67.9	110000	14	ADZ13651_2	ADZ13651_2 Cont	C 224	18.2	64.3	110000	11	ACNA4150_3	ACNA4150_3 Cont
C	152	19	67.9	110000	14	ADZ13620_1	ADZ13620_1 Cont	C 225	18.2	64.3	110000	13	ABD32921_0	ABD32921_0 Mouse can
C	153	19	67.9	110000	14	ADZ13620_2	ADZ13620_2 Cont	C 226	18.2	64.3	142299	10	ADDS0651_1	ADDS0651_1 BAC seque
C	154	19	67.9	144723	11	ACNA4898	ACNA4898 Human gen	C 227	18.2	64.3	142299	14	ADY77909	ADY77909 Human BAC
C	155	19	67.9	276276	11	ACNA4350	ACNA4350 Human gen	C 228	18.2	64.3	142299	14	ADY77909	ADY77909 Human BAC
C	156	19	67.9	276276	11	ADY77909	ADY77909 Human BAC	C 229	18.2	64.3	349981	10	ADDC87619	ADDC87619 Human GPC
C	157	19	67.9	293598	12	ABE33866	ABE33866 Human can	C 230	18.2	64.3	349981	10	ADDC87619	ADDC87619 Human GPC
C	158	18.6	66.4	601	14	ABE33866	ABE33866 Human can	C 231	17.8	63.6	421	4	AAK5167	AAK5167 Human imm
C	159	18.6	66.4	912	5	ABAS8414	ABAS8414 Bacterial	C 232	17.8	63.6	25461	4	ACNA45168	ACNA45168 Drosophi
C	160	18.6	66.4	32184	5	ABAA20589	ABAA20589 Human ner	C 233	17.8	63.6	104514	11	ACNA45168	ACNA45168 Drosophi
C	161	18.6	66.4	38272	4	AAK84952	AAK84952 Human imm	C 234	17.6	62.9	589	5	ABV47896	ABV47896 Human pro
C	162	18.6	66.4	38348	4	AAK84953	AAK84953 Human imm	C 235	17.6	62.9	781	3	AAZ50946	AAZ50946 Photobact
C	163	18.6	66.4	70435	14	ABE32358	ABE32358 Human gen	C 236	17.6	62.9	781	3	AAZ50946	AAZ50946 Photobact
C	164	18.6	66.4	100696	14	ABE32358	ABE32358 Human gen	C 237	17.6	62.9	1190	3	AAZ50947	AAZ50947 NVLUXR fu

C 238	17.6	62.9	1206	3	AAZ50948	Aaz50948 LuxRV fu	311	17.4	62.1	2842	4	ABL28100	ABL28100 Drosophila
C 239	17.6	62.9	1487	13	ADX27604	Adx27604 Plant ful	312	17.4	62.1	3250	3	AAA64292	AAA64292 Nucleocid
C 240	17.6	62.9	1489	13	ADG2673	Adg2673 Transcrip	313	17.4	62.1	3579	3	ABL14708	ABL14708 Drosophila
C 241	17.6	62.9	1504	12	ADM28734	Adm28734 Bacterial	314	17.4	62.1	4643	13	ADT05451	ADT05451 Hemophil
C 242	17.6	62.9	1513	12	ADM28738	Adm28738 Bacterial	315	17.4	62.1	4751	4	ABL16318	ABL16318 Drosophila
C 243	17.6	62.9	1537	5	AAF57764	Aaf57764 Oligonuc	316	17.4	62.1	4816	10	ADD646187	Add646187 Rat gene
C 244	17.6	62.9	1543	14	AED33577	Aed33577 Expressio	317	17.4	62.1	4839	4	ADBE2381	Adbe2381 Drosophila
C 245	17.6	62.9	1545	13	AAA69452	Aaa69452 Plasmid p	318	17.4	62.1	4939	2	ABL14642	Abli14642 Drosophila
C 246	17.6	62.9	1546	5	AAV04373	Aav04373 Sequence	319	17.4	62.1	4993	2	AAO12994	Aaq12994 HincII fr
C 247	17.6	62.9	1546	2	AAF57768	Aaf57768 Oligonuc	320	17.4	62.1	4993	2	AAO12994	Aaq12994 HincII fr
C 248	17.6	62.9	1546	6	ABK24112	Abk24112 Expressio	321	17.4	62.1	5268	4	ABL16316	Abal16316 Drosophila
C 249	17.6	62.9	1546	6	ABL35769	Abli35769 Plasmid p	322	17.4	62.1	5660	5	ABAI9013	Abai9013 Human ner
C 250	17.6	62.9	1546	10	ABT33898	Abt33898 PCFMI656	323	17.4	62.1	6439	13	ADT95948	Adt95948 Drosophila
C 251	17.6	62.9	1546	14	ABE19733	Aeb19733 pMG21 ex	324	17.4	62.1	6439	13	ADT95948	Adt95948 Drosophila
C 252	17.6	62.9	1546	14	AED33581	Aed33581 Expressio	325	17.4	62.1	7529	6	AAI96694	Aai96694 Arabidops
C 253	17.6	62.9	1546	14	ABE19862	Aeb19862 pMG21 ex	326	17.4	62.1	8580	6	AAI96694	Aai96694 Arabidops
C 254	17.6	62.9	1731	13	ADR63920	Adr63920 Cotton CD	327	17.4	62.1	8580	11	ADM39547	Adm39547 DMT polyn
C 255	17.6	62.9	1731	13	ADR63920	Adr63920 Cotton CD	328	17.4	62.1	9128	15	ABE40104	Abe40104 HIV1 RF v
C 256	17.6	62.9	1731	13	ADR63920	Adr63920 Cotton CD	329	17.4	62.1	10478	2	AAE4691	Aae4691 Arabidops
C 257	17.6	62.9	1731	13	ADR63920	Adr63920 Cotton CD	330	17.4	62.1	11453	4	AAE4691	Aae4691 Arabidops
C 258	17.6	62.9	19371	4	AAI05706	Aai05706 Human rep	331	17.4	62.1	11453	4	AAE4691	Aae4691 Arabidops
C 259	17.6	62.9	19371	4	AAI05706	Aai05706 Human rep	332	17.4	62.1	11453	4	AAE4691	Aae4691 Arabidops
C 260	17.6	62.9	19371	4	AAI05706	Aai05706 Human rep	333	17.4	62.1	11453	4	AAE4691	Aae4691 Arabidops
C 261	17.6	62.9	19371	4	AAI05706	Aai05706 Human rep	334	17.4	62.1	11453	4	AAE4691	Aae4691 Arabidops
C 262	17.6	62.9	19371	4	AAI05706	Aai05706 Human rep	335	17.4	62.1	11453	4	AAE4691	Aae4691 Arabidops
C 263	17.6	62.9	19371	4	AAI05706	Aai05706 Human rep	336	17.4	62.1	11453	4	AAE4691	Aae4691 Arabidops
C 264	17.6	62.9	19371	4	AAI05706	Aai05706 Human rep	337	17.4	62.1	11453	4	AAE4691	Aae4691 Arabidops
C 265	17.6	62.9	19371	4	AAI05706	Aai05706 Human rep	338	17.4	62.1	11453	4	AAE4691	Aae4691 Arabidops
C 266	17.6	62.9	19371	4	AAI05706	Aai05706 Human rep	339	17.4	62.1	11453	4	AAE4691	Aae4691 Arabidops
C 267	17.6	62.9	19371	4	AAI05706	Aai05706 Human rep	340	17.4	62.1	11453	4	AAE4691	Aae4691 Arabidops
C 268	17.6	62.9	19371	4	AAI05706	Aai05706 Human rep	341	17.4	62.1	11453	4	AAE4691	Aae4691 Arabidops
C 269	17.6	62.9	19371	4	AAI05706	Aai05706 Human rep	342	17.4	62.1	11453	4	AAE4691	Aae4691 Arabidops
C 270	17.6	62.9	19371	4	AAI05706	Aai05706 Human rep	343	17.4	62.1	11453	4	AAE4691	Aae4691 Arabidops
C 271	17.6	62.9	19371	4	AAI05706	Aai05706 Human rep	344	17.4	62.1	11453	4	AAE4691	Aae4691 Arabidops
C 272	17.6	62.9	19371	4	AAI05706	Aai05706 Human rep	345	17.4	62.1	11453	4	AAE4691	Aae4691 Arabidops
C 273	17.6	62.9	19371	4	AAI05706	Aai05706 Human rep	346	17.4	62.1	11453	4	AAE4691	Aae4691 Arabidops
C 274	17.6	62.9	19371	4	AAI05706	Aai05706 Human rep	347	17.4	62.1	11453	4	AAE4691	Aae4691 Arabidops
C 275	17.6	62.9	19371	4	AAI05706	Aai05706 Human rep	348	17.4	62.1	11453	4	AAE4691	Aae4691 Arabidops
C 276	17.6	62.9	19371	4	AAI05706	Aai05706 Human rep	349	17.4	62.1	11453	4	AAE4691	Aae4691 Arabidops
C 277	17.6	62.9	19371	4	AAI05706	Aai05706 Human rep	350	17.4	62.1	11453	4	AAE4691	Aae4691 Arabidops
C 278	17.6	62.9	19371	4	AAI05706	Aai05706 Human rep	351	17.4	62.1	11453	4	AAE4691	Aae4691 Arabidops
C 279	17.6	62.9	19371	4	AAI05706	Aai05706 Human rep	352	17.4	62.1	11453	4	AAE4691	Aae4691 Arabidops
C 280	17.6	62.9	19371	4	AAI05706	Aai05706 Human rep	353	17.4	62.1	11453	4	AAE4691	Aae4691 Arabidops
C 281	17.6	62.9	19371	4	AAI05706	Aai05706 Human rep	354	17.4	62.1	11453	4	AAE4691	Aae4691 Arabidops
C 282	17.6	62.9	19371	4	AAI05706	Aai05706 Human rep	355	17.4	62.1	11453	4	AAE4691	Aae4691 Arabidops
C 283	17.6	62.9	19371	4	AAI05706	Aai05706 Human rep	356	17.4	62.1	11453	4	AAE4691	Aae4691 Arabidops
C 284	17.6	62.9	19371	4	AAI05706	Aai05706 Human rep	357	17.4	62.1	11453	4	AAE4691	Aae4691 Arabidops
C 285	17.6	62.9	19371	4	AAI05706	Aai05706 Human rep	358	17.4	62.1	11453	4	AAE4691	Aae4691 Arabidops
C 286	17.6	62.9	19371	4	AAI05706	Aai05706 Human rep	359	17.4	62.1	11453	4	AAE4691	Aae4691 Arabidops
C 287	17.6	62.9	19371	4	AAI05706	Aai05706 Human rep	360	17.4	62.1	11453	4	AAE4691	Aae4691 Arabidops
C 288	17.6	62.9	19371	4	AAI05706	Aai05706 Human rep	361	17.4	62.1	11453	4	AAE4691	Aae4691 Arabidops
C 289	17.6	62.9	19371	4	AAI05706	Aai05706 Human rep	362	17.4	62.1	11453	4	AAE4691	Aae4691 Arabidops
C 290	17.6	62.9	19371	4	AAI05706	Aai05706 Human rep	363	17.4	62.1	11453	4	AAE4691	Aae4691 Arabidops
C 291	17.6	62.9	19371	4	AAI05706	Aai05706 Human rep	364	17.4	62.1	11453	4	AAE4691	Aae4691 Arabidops
C 292	17.6	62.9	19371	4	AAI05706	Aai05706 Human rep	365	17.4	62.1	11453	4	AAE4691	Aae4691 Arabidops
C 293	17.6	62.9	19371	4	AAI05706	Aai05706 Human rep	366	17.4	62.1	11453	4	AAE4691	Aae4691 Arabidops
C 294	17.6	62.9	19371	4	AAI05706	Aai05706 Human rep	367	17.4	62.1	11453	4	AAE4691	Aae4691 Arabidops
C 295	17.6	62.9	19371	4	AAI05706	Aai05706 Human rep	368	17.4	62.1	11453	4	AAE4691	Aae4691 Arabidops
C 296	17.6	62.9	19371	4	AAI05706	Aai05706 Human rep	369	17.4	62.1	11453	4	AAE4691	Aae4691 Arabidops
C 297	17.6	62.9	19371	4	AAI05706	Aai05706 Human rep	370	17.4	62.1	11453	4	AAE4691	Aae4691 Arabidops
C 298	17.6	62.9	19371	4	AAI05706	Aai05706 Human rep	371	17.4	62.1	11453	4	AAE4691	Aae4691 Arabidops
C 299	17.6	62.9	19371	4	AAI05706	Aai05706 Human rep	372	17.4	62.1	11453	4	AAE4691	Aae4691 Arabidops
C 300	17.6	62.9	19371	4	AAI05706	Aai05706 Human rep	373	17.4	62.1	11453	4	AAE4691	Aae4691 Arabidops
C 301	17.6	62.9	19371	4	AAI05706	Aai05706 Human rep	374	17.4	62.1	11453	4	AAE4691	Aae4691 Arabidops
C 302	17.6	62.9	19371	4	AAI05706	Aai05706 Human rep	375	17.4	62.1	11453	4	AAE4691	Aae4691 Arabidops
C 303	17.6	62.9	19371	4	AAI05706	Aai05706 Human rep	376	17.4	62.1	11453	4	AAE4691	Aae4691 Arabidops
C 304	17.6	62.9	19371	4	AAI05706	Aai05706 Human rep	377	17.4	62.1	11453	4	AAE4691	Aae4691 Arabidops
C 305	17.6	62.9	19371	4	AAI05706	Aai05706 Human rep	378	17.4	62.1	11453	4	AAE4691	Aae4691 Arabidops
C 306	17.6	62.9	19371	4	AAI05706	Aai05706 Human rep	379	17.4	62.1	11453	4	AAE4691	Aae4691 Arabidops
C 307	17.6	62.9	19371	4	AAI05706	Aai05706 Human rep	380	17.4	62.1	11453	4	AAE4691	Aae4691 Arabidops
C 308	17.6	62.9	19371	4	AAI05706	Aai05706 Human rep	381	17.4	62.1	11453	4	AAE4691	Aae4691 Arabidops
C 309	17.6	62.9	19371	4	AAI05706	Aai05706 Human rep	382	17.4	62.1	11453	4	AAE4691	Aae4691 Arabidops
C 310	17.6	62.9	19371	4	AAI05706	Aai05706 Human rep	383	17.4	62.1	11453	4	AAE4691	Aae4691 Arabidops

C 384	17	60.7	528	4	AAK67947	AAK67947 Human imm	457	17	3800	12	AD035541	Ad035541 Novel mou
C 385	17	60.7	528	4	AAK67946	AAK67946 Human imm	C 458	17	3952	4	ABL02940	Ab102940 Drosophi1
C 386	17	60.7	597	2	AAE68028	AAE68028 H. pylori	C 459	17	4044	4	ABL02942	Ab102942 Drosophi1
C 387	17	60.7	601	13	ADV15561	ADV15561 Human oet	C 460	17	4224	5	AA594229	AA594229 DNA encod
C 388	17	60.7	601	14	ABE33554	ABE33554 Human DNA	C 461	17	4224	5	AA592704	AA592704 DNA encod
C 389	17	60.7	700	4	AAH52940	AAH52940 Human inf	C 462	17	4224	5	AA588559	AA588559 DNA encod
C 390	17	60.7	877	5	ABV22831	ABV22831 Human pro	C 463	17	4350	2	AA212252	AA212252 Netheeria
C 391	17	60.7	877	5	ABV28661	ABV28661 Human pro	C 464	17	4350	8	ACCA1657	ACCA1657 Prokaryot
C 392	17	60.7	894	4	AAH06922	AAH06922 Human CDN	C 465	17	4350	14	ABE94960	ABE94960 N. mening
C 393	17	60.7	921	14	ACG65515	ACG65515 M. xanthu	C 466	17	4350	14	ABE91910	ABE91910 DNA encod
C 394	17	60.7	1083	394	ABX06440	ABX06440 S. pneumo	C 467	17	4355	3	AAZ38917	AAZ38917 Netheeria
C 395	17	60.7	1086	8	ACA49832	ACA49832 Prokaryot	C 468	17	4378	6	ABQ70913	ABQ70913 Listeria
C 396	17	60.7	1086	13	ADK43557	ADK43557 Streptoco	C 469	17	4999	10	AD555113	AD555113 Human gen
C 397	17	60.7	1095	13	ADK93568	ADK93568 Strept S.	C 470	17	5558	2	AAV52226	AAV52226 Streptoco
C 398	17	60.7	1095	14	AEA57438	AEA57438 Strept S.	C 471	17	5573	2	AB144058	AB144058 Drosophi1
C 399	17	60.7	1317	12	AD007229	AD007229 Human org	C 472	17	15572	14	ACI64668	ACI64668 M. xanthu
C 400	17	60.7	1317	14	ABE93649	ABE93649 Human org	C 473	17	2846	6	ADD56096	ADD56096 Mouse Ccn
C 401	17	60.7	1410	6	ABQ68821	ABQ68821 Listeria	C 474	17	2846	9	ADA02458	ADA02458 Mouse Ccn
C 402	17	60.7	1536	6	ABN93116	ABN93116 Staphyloc	C 475	17	2846	10	ADB72197	ADB72197 Mouse Ccn
C 403	17	60.7	1536	13	AD602384	AD602384 Staphyloc	C 476	17	2846	10	AD682929	AD682929 Mouse Ccn
C 404	17	60.7	1536	14	ABE67765	ABE67765 Rice geno	C 477	17	29695	6	ABO78853	ABO78853 Human PPM
C 405	17	60.7	1611	14	AAFP27871	AAFP27871 Human bre	C 478	17	29871	13	ADS36480	ADS36480 Human aut
C 406	17	60.7	1611	9	ACD40273	ACD40273 Human bre	C 479	17	31960	14	ABE44604	ABE44604 Sleep dis
C 407	17	60.7	1611	11	ADM56404	ADM56404 Human CDN	C 480	17	33231	4	ABE61135	ABE61135 Human HPG
C 408	17	60.7	1611	12	AD666766	AD666766 CDNA encd	C 481	17	33231	4	AB114922	AB114922 Drosophi1
C 409	17	60.7	1611	12	AD060276	AD060276 Human NOV	C 482	17	50272	4	AB110156	AB110156 Drosophi1
C 410	17	60.7	1612	12	AD119803	AD119803 Human NOV	C 483	17	53332	6	AA148890	AA148890 Human PFC
C 411	17	60.7	1748	10	ABQ83990	ABQ83990 Mouse uhl	C 484	17	54668	13	ADY87738	ADY87738 Streptoco
C 412	17	60.7	1748	12	ADJ75854	ADJ75854 Marker gen	C 485	17	54668	13	ADY78991	ADY78991 Streptoco
C 413	17	60.7	1813	12	AD008673	AD008673 Clona int	C 486	17	55906	13	ABD33337	ABD33337 Murine ca
C 414	17	60.7	1949	4	AB114059	AB114059 Drosophi1	C 487	17	93001	12	AD079405	AD079405 CENPCL ye
C 415	17	60.7	2006	6	AB215121	AB215121 Arabidops	C 488	17	95001	12	ADH56439	ADH56439 Human hyP
C 416	17	60.7	2006	10	AD5858134	AD5858134 Primary r	C 489	17	96589	9	ADA02708	ADA02708 Human ZFH
C 417	17	60.7	2006	10	AD5858134	AD5858134 Primary r	C 490	17	96589	9	ADB72446	ADB72446 Human ZFH
C 418	17	60.7	2006	13	ADY40942	ADY40942 Rat cardi	C 491	17	97435	10	AD595956	AD595956 Human ZFH
C 419	17	60.7	2006	13	ADY40942	ADY40942 Rat cardi	C 492	17	97435	10	ABD33254	ABD33254 Murine ca
C 420	17	60.7	2351	2	AAZ27259	AAZ27259 Human sec	C 493	17	110000	6	ABN71537	ABN71537 Streptoco
C 421	17	60.7	2351	8	ADA40378	ADA40378 Human sec	C 494	17	110000	8	AA053223_1	AA053223_1 of
C 422	17	60.7	2351	8	ACC50751	ACC50751 Human sec	C 495	17	110000	10	AB566454_06	AB566454_06 Continuation (7 of
C 423	17	60.7	2351	8	AB271414	AB271414 Secretd	C 496	17	110000	13	ADY81204_00	ADY81204_00 Continuation (7 of
C 424	17	60.7	2351	9	ABD91367	ABD91367 Human sec	C 497	17	144392	15	ABP75162	ABP75162 Human pol
C 425	17	60.7	2351	10	ADC73861	ADC73861 Human sec	C 498	17	212231	11	ACN44598	ACN44598 Human gen
C 426	17	60.7	2351	10	ADA56544	ADA56544 Gene encd	C 499	17	242456	14	ABD76150	ABD76150 Mouse CA
C 427	17	60.7	2351	4	AA160312	AA160312 Human pol	C 500	17	312957	13	ADV15290	ADV15290 Human oet
C 428	17	60.7	2488	4	AA158526	AA158526 Human pol	C 501	17	312957	13	ABE32365	ABE32365 Human gen
C 429	17	60.7	2488	5	AD098743	AD098743 DNA encd	C 502	17	312972	13	ADV15293	ADV15293 Human oet
C 430	17	60.7	2486	8	ADBA48503	ADBA48503 Novel hum	C 503	17	312972	14	ABE32394	ABE32394 Human gen
C 431	17	60.7	2486	8	ADA39897	ADA39897 Human sec	C 504	17	313001	14	ADZ70075	ADZ70075 Human inf
C 432	17	60.7	2486	8	ACC50487	ACC50487 Human sec	C 505	17	321019	13	AD536450	AD536450 Human inf
C 433	17	60.7	2486	8	ABZ71269	ABZ71269 Human sec	C 506	17	325093	14	ABE25555	ABE25555 Human ARR
C 434	17	60.7	2486	9	ADB91175	ADB91175 Human sec	C 507	17	329019	13	ABD32707	ABD32707 Human can
C 435	17	60.7	2486	10	ADC73550	ADC73550 Human sec	C 508	16.8	108	6	ABN88421	ABN88421 Mouse gen
C 436	17	60.7	2486	10	ADA56088	ADA56088 Gene encd	C 509	16.8	201	13	ADS38726	ADS38726 Human aut
C 437	17	60.7	2658	13	ADK54630	ADK54630 Plant ful	C 510	16.8	201	13	ADS38511	ADS38511 Human aut
C 438	17	60.7	2733	11	ADL22649	ADL22649 Human dis	C 511	16.8	508	14	AD538724	AD538724 Human aut
C 439	17	60.7	2913	14	AAH14622	AAH14622 Human CDN	C 512	16.8	669	5	AA591869	AA591869 DNA encod
C 440	17	60.7	2913	14	ADZ49660	ADZ49660 Ineuilin s	C 513	16.8	802	5	AD138570	AD138570 Human ova
C 441	17	60.7	2981	4	ADD46639	ADD46639 Human gen	C 514	16.8	802	5	AD138570	AD138570 Human ova
C 442	17	60.7	3113	4	AB118378	AB118378 Drosophi1	C 515	16.8	802	5	AD173440	AD173440 Bacteri1
C 443	17	60.7	3732	4	AB103714	AB103714 Drosophi1	C 516	16.8	845	13	AD611184	AD611184 Mouse spe
C 444	17	60.7	3732	4	AB120934	AB120934 Drosophi1	C 517	16.8	848	9	ADA14465	ADA14465 Bacteri1
C 445	17	60.7	3737	8	AAAD5097	AAAD5097 Mouse Ccn	C 518	16.8	951	11	ACA37101	ACA37101 Rice abio
C 446	17	60.7	3737	8	ADA02459	ADA02459 Mouse Ccn	C 519	16.8	1122	8	ACA37101	ACA37101 Prokaryot
C 447	17	60.7	3737	10	AD872198	AD872198 Mouse Ccn	C 520	16.8	1248	13	ADT16096	ADT16096 Plant CN
C 448	17	60.7	3737	10	AD872198	AD872198 Mouse Ccn	C 521	16.8	1356	10	AD093371	AD093371 E. faeciu
C 449	17	60.7	3751	10	AD857823	AD857823 Toxicity-	C 522	16.8	1371	4	AA552516	AA552516 E. coli D
C 450	17	60.7	3751	10	ABT41704	ABT41704 Toxicity-	C 523	16.8	1371	8	ACA32583	ACA32583 Prokaryot
C 451	17	60.7	3751	11	ADW21750	ADW21750 Rat hepat	C 524	16.8	1530	4	AB117963	AB117963 Drosophi1
C 452	17	60.7	3751	12	AD009496	AD009496 Rat Gi/8-	C 525	16.8	1574	6	ABK34944	ABK34944 Human CDN
C 453	17	60.7	3751	12	ADP72490	ADP72490 Renal tox	C 526	16.8	1616	6	ABO76427	ABO76427 S. cerevi
C 454	17	60.7	3751	12	ADP72270	ADP72270 Renal tox	C 527	16.8	1643	6	AA562515	AA562515 cDNA sequ
C 455	17	60.7	3751	13	ADV40791	ADV40791 Rat cardi	C 528	16.8	1786	14	ABE12000	ABE12000 Tobacco n
C 456	17	60.7	3751	14	ABE11825	ABE11825 Rat surro	C 529	16.8	1786	14	ABE06622	ABE06622 Tobacco n

C 530	16.8	60.0	1786	15	Aef97015 Tobacco n	603	16.6	59.3	47	3	AAZ65823	Aaz65823 Human map
531	16.8	60.0	1806	4	Aaf61036 P. putida	C 604	16.6	59.3	101	2	AAH85681	Aah85681 Human sin
532	16.8	60.0	1815	10	Adg28855 Perennial	C 605	16.6	59.3	121	12	ADK93903	Adk93903 Polynucle
533	16.8	60.0	1873	10	Adg28856 Tail feces	C 606	16.6	59.3	166	3	AAC32807	Aac32807 Human sec
534	16.8	60.0	1905	3	Aac79713 Human sec	C 607	16.6	59.3	197	12	ADP60308	Adp60308 Soybean c
535	16.8	60.0	2133	11	Adm01618 Human cdn	C 608	16.6	59.3	261	12	ADP59352	Adp59352 Soybean c
536	16.8	60.0	2133	14	Aec84548 Human cdn	C 609	16.6	59.3	265	12	ADP60307	Adp60307 Soybean c
537	16.8	60.0	2320	11	Adm03113 Human cdn	C 610	16.6	59.3	294	3	AAA43307	Aaa43307 Xenopus s
C 538	16.8	60.0	2320	14	Aec86043 Human cdn	C 611	16.6	59.3	385	5	ABV03093	Abv03093 Human pro
539	16.8	60.0	2341	10	AdB63479 Human cdn	C 612	16.6	59.3	408	5	ABV12262	Abv12262 Human pro
C 540	16.8	60.0	2464	6	ABQ70708 Listeria	C 613	16.6	59.3	446	5	ABV42330	Abv42330 Human pro
541	16.8	60.0	2647	11	Adm02143 Human cdn	C 614	16.6	59.3	446	5	ABV33407	Abv33407 Human pro
542	16.8	60.0	2793	14	Aec85073 Human cdn	C 615	16.6	59.3	459	6	ABN97022	Abn97022 Gene #352
543	16.8	60.0	2793	5	Aas73600 DNA encod	C 616	16.6	59.3	461	13	ACR90858	Acf90858 Human SIR
544	16.8	60.0	3039	5	Aas82385 DNA encod	C 617	16.6	59.3	552	13	ACR84219	Acf84219 Human SIR
C 545	16.8	60.0	3039	5	Aas77331 DNA encod	C 618	16.6	59.3	644	13	ADK53382	Adk53382 Plant ful
546	16.8	60.0	3039	5	Aas93075 DNA encod	C 619	16.6	59.3	684	8	ACA41274	Aca41274 Prokaryot
547	16.8	60.0	3333	5	Aas89925 DNA encod	C 620	16.6	59.3	684	10	ABZ38326	Abz38326 N. gonorr
548	16.8	60.0	3344	10	ACA54668 Human NF-	C 621	16.6	59.3	784	14	AED14244	Aed14244 Human cum
549	16.8	60.0	3344	14	ADH83145 Human NPK	C 622	16.6	59.3	887	10	ADC24857	Adc24857 Human bre
550	16.8	60.0	3582	14	ABL17962 Drosophi	C 623	16.6	59.3	912	8	ADK29893	Adk29893 DNA encod
C 551	16.8	60.0	3684	11	ADM01821 Human cdn	C 624	16.6	59.3	914	8	ABZ51954	Abz51954 Aspergill
C 552	16.8	60.0	3684	14	AEC84751 Human cdn	C 625	16.6	59.3	1006	15	AER11533	Aer11533 Rape meth
553	16.8	60.0	3910	5	Aas79824 DNA encod	C 626	16.6	59.3	1044	6	ABZ13045	Abz13045 Arabidops
554	16.8	60.0	4092	14	AED81831 Hyperimmu	C 627	16.6	59.3	1125	4	AAH78265	Aah78265 Nucleotid
555	16.8	60.0	4109	3	AAO00316 Human CAM	C 628	16.6	59.3	1439	12	ADP12796	Adp12796 Reference
556	16.8	60.0	4173	8	ABX71051 Novel hum	C 629	16.6	59.3	1544	3	AAC52745	Aac52745 Arabidops
557	16.8	60.0	4275	4	ABL09776 Drosophi	C 630	16.6	59.3	1750	2	AAT83352	Aat83352 B. c. alka
558	16.8	60.0	4277	4	ABL11776 Drosophi	C 631	16.6	59.3	1764	13	ADG46812	Adg46812 Bacteri
559	16.8	60.0	4845	12	ADQ23064 Human sof	C 632	16.6	59.3	1765	13	ADT05175	Adt05175 Hemophil
560	16.8	60.0	4845	12	ADQ25056 Human sof	C 633	16.6	59.3	2018	13	ADK54881	Adk54881 Plant ful
C 561	16.8	60.0	5008	8	ABK63351 Human sof	C 634	16.6	59.3	2188	2	AAQ79726	Aaq79726 Mouse L5/
C 562	16.8	60.0	6347	14	AED11934 Tobacco n	C 635	16.6	59.3	2188	2	AAT62441	Aat62441 Mouse gro
C 563	16.8	60.0	6347	14	AEE06617 Tobacco n	C 636	16.6	59.3	2213	12	AAV07384	Aav07384 DNA encod
C 564	16.8	60.0	6347	15	AEE97010 Tobacco n	C 637	16.6	59.3	2216	2	AAV72084	Aav72084 Mouse MSP
565	16.8	60.0	8675	12	ADQ97649 Mouse can	C 638	16.6	59.3	2324	4	ABL19952	Ab119952 Drosophi
566	16.8	60.0	11574	4	AA546244 DNA encod	C 639	16.6	59.3	2767	11	ADM64616	Adm64616 Mose (Phy
567	16.8	60.0	12822	9	ABK42692 Genomic s	C 640	16.6	59.3	3276	13	ADK60882	Adk60882 Plant ful
568	16.8	60.0	12822	9	ADB60848 Connectio	C 641	16.6	59.3	3336	4	AA555220	Aas55220 Stephyloc
569	16.8	60.0	14226	6	ABL42494 Bacteriop	C 642	16.6	59.3	3389	10	ADK07125	Adk07125 Novel cod
570	16.8	60.0	17369	4	ABL02754 Drosophi	C 643	16.6	59.3	3389	14	ADU40282	Adu40282 Novel hum
C 571	16.8	60.0	29693	14	AEE04799 Cancer-ab	C 644	16.6	59.3	3450	8	ACC62270	Acc62270 Human NOV
572	16.8	60.0	57367	14	AEB39164 L. pneumo	C 645	16.6	59.3	3534	4	AAH78264	Aah78264 Nucleotid
C 573	16.8	60.0	77425	6	ABK83502 Human cdn	C 646	16.6	59.3	3534	6	ABQ74273	Abq74273 Human 441
574	16.8	60.0	80275	12	ADQ97310 Mouse can	C 647	16.6	59.3	3534	10	ADJ37494	Adj37494 Human tra
575	16.8	60.0	88607	12	ADQ97648 Mouse can	C 648	16.6	59.3	3534	12	ADJ27973	Adj27973 Pathogen
C 576	16.8	60.0	105295	14	AEB35712 L. pneumo	C 649	16.6	59.3	3822	8	ABT14927	Abt14927 Stephyloc
577	16.8	60.0	110000	4	AAI99683_41	C 650	16.6	59.3	3822	8	ACF72769	Acf72769 Stephyloc
578	16.8	60.0	110000	4	AAI99683_41	C 651	16.6	59.3	3825	8	ACA19783	Aca19783 Prokaryot
C 579	16.8	60.0	110000	6	ABA90521_14	C 652	16.6	59.3	3855	13	ADR44057	Adr44057 Pasteurel
580	16.8	60.0	110000	10	ABO84281_0	C 653	16.6	59.3	3858	13	ADR44058	Adr44058 Pasteurel
581	16.8	60.0	110000	11	ACN44932_2	C 654	16.6	59.3	4245	6	AAL44686	Aal44686 Human tra
582	16.8	60.0	110000	12	ADQ97050_4	C 655	16.6	59.3	4380	15	AAH91839	Aah91839 Pasteurel
583	16.8	60.0	110000	14	AEB39172_2	C 656	16.6	59.3	4380	15	AEE70376	Aee70376 Pasteurel
584	16.8	60.0	110000	14	AEB39175_30	C 657	16.6	59.3	4381	2	AAO10519	Aao10519 Pasteurel
585	16.8	60.0	110000	14	AEB39175_32	C 658	16.6	59.3	4417	9	AAL62510	Aal62510 Human tra
586	16.8	60.0	110000	14	AEB42401_28	C 659	16.6	59.3	4457	3	AAH73385	Aah73385 Heterosig
587	16.8	60.0	110000	14	AEB42401_30	C 660	16.6	59.3	4522	2	AAV74350	Aav74350 Stephyloc
588	16.8	60.0	110000	14	AEB42401_15	C 661	16.6	59.3	6751	2	AAQ79727	Aaq79727 Mouse L5/
589	16.8	60.0	110000	14	AEB42736_4	C 662	16.6	59.3	6751	2	AAT62442	Aat62442 Mouse gro
590	16.8	60.0	149671	6	ABK84797 Human cdn	C 663	16.6	59.3	7204	10	ABR411748	Abt411748 Toxicity
591	16.8	60.0	149671	9	ADB70361 Mosein cd	C 664	16.6	59.3	7204	13	ADP72599	Adp72599 Renal tox
592	16.8	60.0	149671	12	ADJ377140 Human mal	C 665	16.6	59.3	7204	13	ADR14992	Adr14992 Ret elect
593	16.8	60.0	165156	13	ADK36459 Human aut	C 666	16.6	59.3	7221	6	ABQ74272	Abq74272 Human 441
C 594	16.8	60.0	175603	13	ADQ97554 Mouse can	C 667	16.6	59.3	7221	12	ADD337492	Add337492 Human tra
595	16.8	60.0	229301	14	AEB35719 L. pneumo	C 668	16.6	59.3	7221	12	ADJ27971	Adj27971 Human 441
C 596	16.8	60.0	260027	11	ACN44046 Human gen	C 669	16.6	59.3	7232	4	AAK79332	Aak79332 Human imm
597	16.8	60.0	349881	10	ADK66642 Human GPC	C 670	16.6	59.3	7232	4	AAK72237	Aak72237 Human imm
C 598	16.8	60.0	349881	10	ADK66642 Human GPC	C 671	16.6	59.3	7277	4	AAH78265	Aah78265 Nucleotid
599	16.6	59.3	41	2	AAV50720 Brassaica	C 672	16.6	59.3	7347	13	ADG97793	Adg97793 Protein E
600	16.6	59.3	41	2	AAV50729 Brassaica	C 673	16.6	59.3	7347	13	ADU20792	Adu20792 Human ATP
601	16.6	59.3	41	2	AAV50728 Brassaica	C 674	16.6	59.3	7355	8	AAU49507	Aau49507 Human TRI
602	16.6	59.3	41	2	AAV50726 Brassaica	C 675	16.6	59.3	7683	13	ADR84477	Adr84477 Aspergill

C 676	16.6	59.3	7943	4	AAK85948	Aak85948 Human imm	C 749	16.4	58.6	1075	4	AAH32611	Aah32611 Human sec
C 677	16.6	59.3	7968	10	ADCI9779	Adci9779 H. influenzae	C 750	16.4	58.6	1092	2	AAZ19223	Aaz19223 M. tuberculosis
C 678	16.6	59.3	8217	14	ADM94175	Adm94175 Staphyloc	C 751	16.4	58.6	1092	2	AAZ19223	Aaz19223 M. tuberculosis
C 679	16.6	59.3	10214	14	ADM94174	Adm94174 Staphyloc	C 752	16.4	58.6	1104	3	AAK47642	Aak47642 Arabidops
C 680	16.6	59.3	10630	4	ABL05498	Abi05498 Drosophila	C 753	16.4	58.6	1158	10	ACF70585	Actf70585 Prochlorb
C 681	16.6	59.3	11244	4	AAK74606	Aak74606 Human imm	C 754	16.4	58.6	1167	13	ADT44282	Adt44282 Bacterial
C 682	16.6	59.3	15236	4	AAK74605	Aak74605 Human imm	C 755	16.4	58.6	1201	3	AAK49054	Aak49054 Arabidops
C 683	16.6	59.3	26811	2	AAK20253	Aak20253 Borrelia	C 756	16.4	58.6	1224	4	ABL21641	Abi21641 Drosophila
C 684	16.6	59.3	41540	11	ACN44338	Acn44338 Mouse gen	C 757	16.4	58.6	1229	2	AAQ46140	Aaq46140 NKX1.3/2
C 685	16.6	59.3	95484	12	ADQ97298	Adq97298 Mouse can	C 758	16.4	58.6	1272	8	ACA53296	Acas5296 Prokaryot
C 686	16.6	59.3	110000	6	ABO6245_09	Continuation (10 o	C 759	16.4	58.6	1281	11	ACM92123	Acme92123 Breast ca
C 687	16.6	59.3	110000	6	ABQ6719_03	Continuation (9 of	C 760	16.4	58.6	1291	12	ADL12859	Adl12859 Human sec
C 688	16.6	59.3	110000	14	AEBA39175_03	Continuation (4 of	C 761	16.4	58.6	1310	3	AAZ34554	Aaz34554 Human sec
C 689	16.6	59.3	110000	14	AEBA42401_02	Continuation (3 of	C 762	16.4	58.6	1311	10	ADR60103	Adrk60103 Plant DNA
C 690	16.6	59.3	110000	14	AEBA42401_03	Continuation (4 of	C 763	16.4	58.6	1339	3	AAK48761	Aak48761 Arabidops
C 691	16.6	59.3	110000	14	AEBA42737_13	Continuation (14 o	C 764	16.4	58.6	1382	5	AAK74298	Aak74298 DNA encod
C 692	16.6	59.3	110000	14	AEBA42736_05	Continuation (6 of	C 765	16.4	58.6	1439	3	AAK40284	Aak40284 Arabidops
C 693	16.6	59.3	110000	15	AEBA42737_09	Continuation (9 of	C 766	16.4	58.6	1492	10	ADCC30373	Adcc30373 Human nov
C 694	16.6	59.3	110000	15	AEBA42737_09	Continuation (9 of	C 767	16.4	58.6	1649	2	AAK13356	Aak13356 Enterococ
C 695	16.6	59.3	118584	10	ADCE87623	Adce87623 Human gpc	C 768	16.4	58.6	1649	6	ABK39151	Abk39151 Enterococ
C 696	16.6	59.3	126932	15	AEBA51725	Aeba51725 Human rib	C 769	16.4	58.6	1724	4	AAH32609	Aah32609 Human sec
C 697	16.6	59.3	194588	14	AEBA51725	Aeba51725 Human rib	C 770	16.4	58.6	1761	4	AAK68428	Aak68428 Human imm
C 698	16.6	59.3	200000	12	ADCA7191	Adca7191 DNA seque	C 771	16.4	58.6	1761	4	AAK68427	Aak68427 Human imm
C 699	16.6	59.3	200622	14	AEBA39167	Aeba39167 L. pneumo	C 772	16.4	58.6	1789	13	ADK48649	Adk48649 Bacterial
C 700	16.6	59.3	207461	14	AEBA39170	Aeba39170 L. pneumo	C 773	16.4	58.6	1860	4	AAK57030	Aak57030 C. pneumo
C 701	16.6	59.3	349980	13	ADT057237	Adt057237 Haemophil	C 774	16.4	58.6	1860	6	ABL91273	Abi91273 Chlamydia
C 702	16.6	58.6	121	10	ADH92425	Adh92425 Human gen	C 775	16.4	58.6	1860	10	ADDA2824	Adda2824 Chlamydia
C 703	16.4	58.6	184	2	AAO85523	Aao85523 DNA probe	C 776	16.4	58.6	1875	8	ABV75001	Abv75001 C. pneumo
C 704	16.4	58.6	185	2	AAO85532	Aao85532 DNA probe	C 777	16.4	58.6	2005	4	AAK81805	Aak81805 Human sec
C 705	16.4	58.6	210	12	ACH82137	Ach82137 Human gen	C 778	16.4	58.6	2005	4	AAK81805	Aak81805 Human sec
C 706	16.4	58.6	216	8	ABX43084	Abx43084 Bovine ES	C 779	16.4	58.6	2060	6	AAK81805	Aak81805 Human sec
C 707	16.4	58.6	257	13	ADV39958	Adv39958 Rat card1	C 780	16.4	58.6	2097	13	ADT45518	Adt45518 Human sec
C 708	16.4	58.6	295	2	AAO61051	Aao61051 Human bra	C 781	16.4	58.6	2147	14	ABE11429	Abel11429 Bacterial
C 709	16.4	58.6	300	3	AAO00946	Aao00946 Human col	C 782	16.4	58.6	2157	6	ABZ66831	Abz66831 Arabidops
C 710	16.4	58.6	354	6	ABZ08713	Abz08713 Human lei	C 783	16.4	58.6	2157	6	ABZ66831	Abz66831 Arabidops
C 711	16.4	58.6	400	8	ABX37785	Abx37785 Bovine ES	C 784	16.4	58.6	2410	6	ABZ66831	Abz66831 Arabidops
C 712	16.4	58.6	414	8	ABX43970	Abx43970 Bovine ES	C 785	16.4	58.6	2417	4	AAH15011	Aah15011 Human CDN
C 713	16.4	58.6	430	8	ABX62051	Abx62051 Arabidops	C 786	16.4	58.6	2518	4	AAH46774	Aah46774 Human sec
C 714	16.4	58.6	504	11	ADT96410	Adt96410 Colon can	C 787	16.4	58.6	2522	6	ABA96879	Abas96879 Retinal b
C 715	16.4	58.6	504	11	ADT96410	Adt96410 Colon can	C 788	16.4	58.6	2640	11	ACL27116	Actl27116 Rice abio
C 716	16.4	58.6	509	4	AAK57741	Aak57741 Human imm	C 789	16.4	58.6	2679	10	ADK63014	Adk63014 Human CDN
C 717	16.4	58.6	514	12	ACH68436	Ach68436 Human gen	C 790	16.4	58.6	2684	10	ACF588261	Actf588261 ENSR00000
C 718	16.4	58.6	544	6	ABK45511	Abk45511 cDNA encod	C 791	16.4	58.6	2684	10	ADK63014	Adk63014 Human CDN
C 719	16.4	58.6	581	13	ADRS9380	Adrs9380 Cotton cd	C 792	16.4	58.6	2684	12	ADN06056	Adn06056 AntLipocor
C 720	16.4	58.6	581	13	ACNS3030	Acns3030 Cotton an	C 793	16.4	58.6	2684	12	ADN06056	Adn06056 AntLipocor
C 721	16.4	58.6	582	13	ADRS60447	Adrs60447 Cotton cd	C 794	16.4	58.6	3007	9	ACF35894	Actf35894 Human ARP
C 722	16.4	58.6	601	13	ADV15466	Adv15466 Human oet	C 795	16.4	58.6	3007	12	ADQ74862	Adq74862 Human and
C 723	16.4	58.6	601	13	ADV15465	Adv15465 Human oet	C 796	16.4	58.6	3015	6	ABZ66839	Abz66839 Arabidops
C 724	16.4	58.6	601	13	ADV15739	Adv15739 Human oet	C 797	16.4	58.6	3080	12	ADCC35986	Adcc35986 Novel mou
C 725	16.4	58.6	601	13	ADV15740	Adv15740 Human oet	C 798	16.4	58.6	3110	3	AAA10672	Aaa10672 Human Bm
C 726	16.4	58.6	601	14	AEBA32490	Aeba32490 Human DNA	C 799	16.4	58.6	3110	12	ADP88598	Adp88598 Human POV
C 727	16.4	58.6	601	14	AEBA32490	Aeba32490 Human DNA	C 800	16.4	58.6	3110	12	ADZ49076	Adz49076 Insulin s
C 728	16.4	58.6	601	14	AEBA32491	Aeba32491 Human DNA	C 801	16.4	58.6	3215	14	ADM16810	Adm16810 Hepatic tis
C 729	16.4	58.6	601	14	AEBA32491	Aeba32491 Human DNA	C 802	16.4	58.6	3351	13	ADT44214	Adt44214 Bacterial
C 730	16.4	58.6	612	10	ADCC92687	Adcc92687 E. faeciu	C 803	16.4	58.6	4788	4	ABL12772	Abi12772 Drosophila
C 731	16.4	58.6	615	13	ADQ79543	Adq79543 Novel can	C 804	16.4	58.6	4788	4	ABL12772	Abi12772 Drosophila
C 732	16.4	58.6	622	5	ABV52155	Abv52155 Human pro	C 805	16.4	58.6	4961	5	AAH45501	Aah45501 Drosophila
C 733	16.4	58.6	625	13	ADQ52294	Adq52294 Human can	C 806	16.4	58.6	5033	8	ABZ74587	Abz74587 secreted
C 734	16.4	58.6	628	13	ACNS3917	Acns3917 Cotton an	C 807	16.4	58.6	5033	10	ADCC21005	Adcc21005 Human sec
C 735	16.4	58.6	718	5	AAST77231	Aast77231 DNA encod	C 808	16.4	58.6	5033	10	ABZ68115	Abz68115 Human sec
C 736	16.4	58.6	718	5	AAST77231	Aast77231 DNA encod	C 809	16.4	58.6	5241	12	ADQ40492	Adq40492 Plant CDN
C 737	16.4	58.6	721	3	AAAC38816	Aaac38816 Arabidops	C 810	16.4	58.6	5241	12	ADQ40492	Adq40492 Plant CDN
C 738	16.4	58.6	723	10	ADCC08043	Adcc08043 Rice DNA	C 811	16.4	58.6	6236	4	AAK41724	Aak41724 Genomic s
C 739	16.4	58.6	753	11	ACLI26296	Accli26296 Rice abio	C 812	16.4	58.6	6236	4	AAK41724	Aak41724 Genomic s
C 740	16.4	58.6	756	10	ADH83549	Adh83549 Enterococ	C 813	16.4	58.6	6374	4	ABL09964	Abi09964 Drosophila
C 741	16.4	58.6	759	10	ADH83549	Adh83549 Enterococ	C 814	16.4	58.6	6645	8	ABX34714	Abx34714 Human mdd
C 742	16.4	58.6	759	11	ACLI0596	Accli0596 Rice abio	C 815	16.4	58.6	7966	4	ABL19154	Abi19154 Drosophila
C 743	16.4	58.6	769	11	ADMA5610	Adma5610 Insect re	C 816	16.4	58.6	8586	4	ABL07988	Abi07988 Drosophila
C 744	16.4	58.6	878	4	AAK67270	Aak67270 Human imm	C 817	16.4	58.6	8682	4	ABL10382	Abi10382 Drosophila
C 745	16.4	58.6	931	6	ABN98614	Abn98614 Arabidops	C 818	16.4	58.6	9031	10	ADG64990	Adg64990 Open read
C 746	16.4	58.6	999	13	ADT20083	Adt20083 Plant CDN	C 819	16.4	58.6	9835	4	ABL21524	Abi21524 Drosophila
C 747	16.4	58.6	1004	3	AACT5110	Aact5110 Arabidops	C 820	16.4	58.6	9903	4	ABL09916	Abi09916 Drosophila
C 748	16.4	58.6	1006	3	AACT34746	Aact34746 Arabidops	C 821	16.4	58.6	10478	2	AAV84691	Aav84691 Arabidops

822	16.4	58.6	12060	4	AAK80004	AAK80004 Human imm	895	16.2	57.9	41	2	AAV50725	AAV50725 Brassica
823	16.4	58.6	12687	2	AAQ98814	AAQ98814 EAV GL ge	896	16.2	57.9	41	2	AAV50727	AAV50727 Brassica
824	16.4	58.6	12687	2	AAQ98814	AAQ98814 Equine ar	897	16.2	57.9	57	2	AAV18144	AAV18144 Oligo use
825	16.4	58.6	12704	10	ADG64999	ADG64999 Equine ar	898	16.2	57.9	80	2	AAV73211	AAV73211 C. utilis
826	16.4	58.6	12827	2	AAV09036	AAV09036 Equine ar	899	16.2	57.9	124	2	AAQ13879	AAQ13879 Gp120 maj
827	16.4	58.6	13203	4	ABL30124	ABL30124 Drosophill	900	16.2	57.9	151	8	ACC79386	ACC79386 Mouse DST
828	16.4	58.6	15528	2	AAV09039	AAV09039 Equine ar	901	16.2	57.9	183	6	ABL61726	ABL61726 Colton ade
829	16.4	58.6	22659	14	AEBO4761	AEBO4761 Cancer-as	902	16.2	57.9	265	2	AAV73186	AAV73186 C. utilis
830	16.4	58.6	28180	4	AAK68939	AAK68939 Human imm	903	16.2	57.9	292	3	AAAC21393	AAAC21393 Human sec
831	16.4	58.6	28344	4	ABL21640	ABL21640 Drosophill	904	16.2	57.9	303	13	ACF833397	ACF833397 Human SIR
832	16.4	58.6	28564	10	ADD48832	ADD48832 Human gen	905	16.2	57.9	339	14	ACL65544	ACL65544 M. xanthu
833	16.4	58.6	32328	10	ADDA7694	ADDA7694 Human gen	906	16.2	57.9	363	10	ADH83831	ADH83831 Enterococ
834	16.4	58.6	44324	4	ABL09994	ABL09994 Drosophill	907	16.2	57.9	426	10	ADH83920	ADH83920 Bacterial
835	16.4	58.6	44324	11	ACNA84724	ACNA84724 Human imm	908	16.2	57.9	471	8	ACA45946	ACA45946 Prokaryot
836	16.4	58.6	48727	4	AAK67375	AAK67375 Mouse gen	909	16.2	57.9	509	12	ADPE6125	ADPE6125 Human CDN
837	16.4	58.6	55829	13	ABD33512	ABD33512 Human can	910	16.2	57.9	534	10	ABT40874	ABT40874 Toxicity
838	16.4	58.6	73544	13	ABBS56296	ABBS56296 Human tra	911	16.2	57.9	534	13	ADV39906	ADV39906 Rat card1
839	16.4	58.6	86149	12	ADQ97278	ADQ97278 Human can	912	16.2	57.9	566	6	ABQ46314	ABQ46314 Oligonuc1
840	16.4	58.6	88441	12	ADQ97097	ADQ97097 Mouse can	913	16.2	57.9	566	6	ABQ46315	ABQ46315 Oligonuc1
841	16.4	58.6	89736	13	ABD22966_10	Continuation (11 o	914	16.2	57.9	577	5	ABV55965	ABV55965 Human pro
842	16.4	58.6	89736	13	ABD22966_10	Continuation (11 o	915	16.2	57.9	582	3	AAAS8292	AAAS8292 DNA encod
843	16.4	58.6	96596	9	ADA02504	Ada02504 Human BAC	916	16.2	57.9	612	6	ABLS5999	ABLS5999 Human gro
844	16.4	58.6	96596	10	ADBT2242	ADBT2242 Human BAC	917	16.2	57.9	637	6	ABQ18069	ABQ18069 Oligonuc1
845	16.4	58.6	96596	10	ADBS5752	ADBS5752 Human BAC	918	16.2	57.9	637	6	ABQ18068	ABQ18068 Oligonuc1
846	16.4	58.6	99544	14	ADZ13273	ADZ13273 Human can	919	16.2	57.9	648	10	ADC90952	ADC90952 E. faeciu
847	16.4	58.6	99588	11	ACNA5034	ACNA5034 Human gen	920	16.2	57.9	692	13	ADX37150	ADX37150 Plant ful
848	16.4	58.6	110000	2	AAV21209_12	Continuation (5 of	921	16.2	57.9	694	13	ADX9457	ADX9457 Plant ful
849	16.4	58.6	110000	2	AAV21209_12	Continuation (13 o	922	16.2	57.9	788	14	ACL63584	ACL63584 M. xanthu
850	16.4	58.6	110000	2	AAV21209_12	Continuation (13 o	923	16.2	57.9	810	4	AAI96065	AAI96065 Human neu
851	16.4	58.6	110000	6	ABO69245_21	Continuation (12 o	924	16.2	57.9	816	6	ABQ45906	ABQ45906 Oligonuc1
852	16.4	58.6	110000	6	ABO69245_21	Continuation (22 o	925	16.2	57.9	816	6	ABQ45907	ABQ45907 Oligonuc1
853	16.4	58.6	110000	6	ABO69245_22	Continuation (23 o	926	16.2	57.9	845	13	ADX12504	ADX12504 Plant ful
854	16.4	58.6	110000	8	ABX16390_2	Continuation (3 of	927	16.2	57.9	866	13	ADX36971	ADX36971 Plant ful
855	16.4	58.6	110000	8	ABX16390_2	Continuation (3 of	928	16.2	57.9	927	2	AAV73179	AAV73179 C. utilis
856	16.4	58.6	110000	9	ADAI3411_0	ADAI3411 Human tra	929	16.2	57.9	975	15	AEF11545	AEF11545 Rape meth
857	16.4	58.6	110000	10	ACPF6385_5	ACPF6385 Human chr	930	16.2	57.9	1000	14	AEBS8659	AEBS8659 Human pho
858	16.4	58.6	110000	10	ACPF67367_34	Continuation (35 o	931	16.2	57.9	1038	12	ADL03332	ADL03332 DNA encod
859	16.4	58.6	110000	10	ACPF67367_36	Continuation (37 o	932	16.2	57.9	1107	14	ABE67176	ABE67176 Rice geno
860	16.4	58.6	110000	10	ACPF65386_11	Continuation (12 o	933	16.2	57.9	1216	14	ADZ61296	ADZ61296
861	16.4	58.6	110000	11	ACNA5090_1	Continuation (2 of	934	16.2	57.9	1425	5	AAAS84262	AAAS84262
862	16.4	58.6	110000	12	ADQ59446_0	Adq59446 Human can	935	16.2	57.9	1434	12	ADL14280	ADL14280 Plant tra
863	16.4	58.6	110000	13	ABD32966_08	Continuation (9 of	936	16.2	57.9	1434	12	ADL14280	ADL14280 Plant tra
864	16.4	58.6	110000	14	ADZ13631_3	Continuation (4 of	937	16.2	57.9	1563	8	ACA33748	ACA33748 Prokaryot
865	16.4	58.6	110000	14	ADZ13620_3	Continuation (4 of	938	16.2	57.9	1563	8	ACA33748	ACA33748 Prokaryot
866	16.4	58.6	110000	14	ADZ13757_0	Adz13757 Human can	939	16.2	57.9	1616	13	ADXS4113	ADXS4113 Plant ful
867	16.4	58.6	12767	13	ABD32657	ABD32657 Mouse can	940	16.2	57.9	2026	13	ADT16798	ADT16798 Plant CDN
868	16.4	58.6	137000	12	ADH77370	ADH77370 Human PTP	941	16.2	57.9	2156	9	AAH45046	AAH45046
869	16.4	58.6	137000	15	AEBS6219	AEBS6219 Human PTP	942	16.2	57.9	2156	9	ADBL1228	ADBL1228
870	16.4	58.6	139308	8	ADBI2769	ADBI2769 Human PRX	943	16.2	57.9	2199	10	ADC08341	ADC08341 Rice DNA
871	16.4	58.6	153170	12	ADQ17382	Adq17382 Human sof	944	16.2	57.9	2251	2	AAAS0895	AAAS0895 Soybean p
872	16.4	58.6	160482	11	ACNA3914	ACNA3914 Human gen	945	16.2	57.9	2358	2	AAV45134	AAV45134 Mouse Fas
873	16.4	58.6	164991	10	ADL13635	ADL13635 Oetecarth	946	16.2	57.9	2360	4	AAV72909	AAV72909 Mouse dax
874	16.4	58.6	168828	14	ADZ13592	Adz13592 Human can	947	16.2	57.9	2521	10	ADD68418	ADD68418 Human smo
875	16.4	58.6	171398	14	ADZ13359	Adz13359 Human can	948	16.2	57.9	2653	13	ADX49853	ADX49853 Plant ful
876	16.4	58.6	176870	10	ADL13512	Adl13512 Oetecarth	949	16.2	57.9	2919	13	ADMS47919	ADMS47919 Bacterial
877	16.4	58.6	191584	13	ABD33586	ABD33586 Human can	950	16.2	57.9	2970	14	AEBC4628	AEBC4628 Human CDN
878	16.4	58.6	191584	13	ABD33586	ABD33586 Human can	951	16.2	57.9	2970	14	AEBC4628	AEBC4628 Human CDN
879	16.4	58.6	199994	14	AEAO7496	AEAO7496 Human can	952	16.2	57.9	2982	12	ACOA24241	ACOA24241 Prokaryot
880	16.4	58.6	200000	12	ADOA47190	ADOA47190 DNA Seqe	953	16.2	57.9	3324	8	ADCA2892	ADCA2892 4 synthe
881	16.4	58.6	213251	6	ABO67193	ABO67193 Listeria	954	16.2	57.9	3858	4	ABL16114	ABL16114 Drosophill
882	16.4	58.6	215974	12	ADQ97523	ADQ97523 Human can	955	16.2	57.9	3858	4	ADBE11731	ADBE11731 Human gen
883	16.4	58.6	309400	5	AAH68534	AAH68534 C glutami	956	16.2	57.9	4682	9	ADBE11735	ADBE11735 Human gen
884	16.4	58.6	312957	13	ADVI5290	ADVI5290 Human oet	957	16.2	57.9	4682	9	ADBI17230	ADBI17230 Oryzias
885	16.4	58.6	312957	14	AEBS2365	AEBS2365 Human oet	958	16.2	57.9	5255	13	ADRO7597	ADRO7597 Full leng
886	16.4	58.6	312972	13	ADVI5293	ADVI5293 Human oet	959	16.2	57.9	5815	6	ABL33234	ABL33234 Human imm
887	16.4	58.6	312972	14	AEBS2394	AEBS2394 Human oet	960	16.2	57.9	6032	6	AAAS63310	AAAS63310 Chemical1
888	16.4	58.6	313001	14	ADZ70075	Adz70075 Human ins	961	16.2	57.9	6247	6	ABL32275	ABL32275 Human imm
889	16.4	58.6	313287	13	ABD33100	ABD33100 Human can	962	16.2	57.9	6283	6	ABK39991	ABK39991 Human che
890	16.4	58.6	32019	13	ADBS36450	ADBS36450 Human aut	963	16.2	57.9	6283	6	ABL32834	ABL32834 Human imm
891	16.4	58.6	325019	13	ABD32707	ABD32707 Human can	964	16.2	57.9	6858	15	AEBS5363	AEBS5363 Laminaria
892	16.4	58.6	349980	5	AAH68526	AAH68526 C glutami	965	16.2	57.9	9619	2	AAI31134	AAI31134 Enterococ
893	16.4	58.6	349980	5	AAH68533	AAH68533 C glutami	966	16.2	57.9	9619	6	ABBS98929	ABBS98929 Enterococ
894	16.4	58.6	349980	5	AAH68532	AAH68532 C glutami	967	16.2	57.9	22885	11	ACNA4576	ACNA4576 Mouse gen

968	16.2	57.9	32872	4	ABL04072	Abi04072 Drosophila
C 969	16.2	57.9	107543	13	ABD33524	Abd33524 Human can
C 970	16.2	57.9	110000	6	ABX08336_03	Continuation (4 of
C 971	16.2	57.9	110000	12	ADJ25985_03	Continuation (4 of
C 972	16.2	57.9	110000	12	ADN97989_03	Continuation (4 of
C 973	16.2	57.9	110000	12	ADOS0281_03	Continuation (4 of
C 974	16.2	57.9	110000	14	AE885185_03	Continuation (4 of
C 975	16.2	57.9	168325	11	ACN44484	Acn44484 Mouse gen
C 976	16.2	57.9	260160	4	ADQ20017	Adq20017 Human sof
C 977	16.2	57.9	269223	12	AAF28554	Aaf28554 Genomic f
C 978	16	57.1	70	2	AAT71528	Aat71528 Glioblast
C 979	16	57.1	149	6	ABL86945	Abi86945 Human ova
C 980	16	57.1	149	14	ACL62395	ACL62395 Human col
C 981	16	57.1	149	14	ACL54642	ACL54642 Human bre
C 982	16	57.1	171	4	AAL24284	Aal24284 Human col
C 983	16	57.1	243	12	ADP28157	Adp28157 Human sec
C 984	16	57.1	272	4	AAS58298	Aas58298 cDNA #974
C 985	16	57.1	285	4	AAS59455	Aas59455 Streptoco
C 986	16	57.1	285	13	ADR44507	Adr44507 Streptoco
C 987	16	57.1	293	6	ABK31067	Abk31067 Plant dwa
C 988	16	57.1	351	12	ADL02689	Adl02689 DNA encod
C 989	16	57.1	379	4	AAI82699	Aai82699 Human pol
C 990	16	57.1	399	13	ADOS2446	Ados2446 Novel can
C 991	16	57.1	405	4	AAF59800	Aaf59800 Moraxella
C 992	16	57.1	442	12	ADQ17797	Adq17797 Human sof
C 993	16	57.1	444	3	AAC74671	Aac74671 Human ORF
C 994	16	57.1	488	9	ACH32021	Ach32021 Human end
C 995	16	57.1	503	9	ACH37381	Ach37381 Human end
C 996	16	57.1	508	6	ABO44663	Abq44663 Oligonuc
C 997	16	57.1	508	6	ABO44662	Abq44662 Oligonuc
C 998	16	57.1	509	11	ACN87951	Acn87951 Breast ca
C 999	16	57.1	519	13	ADR63921	Adr63921 Cotton CD
1000	16	57.1	519	13	ACNS9775	Acns9775 Cotton gy

ALIGNMENTS

RESULT 1
AAT03309 standard; DNA; 28 BP.

AC AAT03309;
27-AUG-2003 (revised)
11-APR-1996 (first entry)

DE HPV E6 gene contg. plasmid restriction fragment hybridisation probe.
XX Enzyme inhibition suppressor; polymerase; detergent; SDS; LLS; probe;
KM sodium dodecyl sulphate; lithium lauryl sulphate; detection;
KM polymerase chain reaction; ss.

XX Human papillomavirus.
OS BP671473-A1.
XX 13-SEP-1995.
PD 10-MAR-1995; 95BP-00103520.
PF 10-MAR-1994; 94US-00212131.
PR 10-MAR-1994; 94US-00212131.
XX (GENP-) GEN-PROBE INC.
PA Kacian DL, Mcallister DL;
XX WPI; 1995-312773/41.
DR Use of nonionic detergent to suppress enzyme inhibition by ionic
XX detergent - partic. to enhance activity of e.g. DNA or RNA polymerase in
PT nucleic acid amplification(s).
XX

PS Example 4; Page 9; 20pp; English.
XX AAT03307-08 and T3309 are a pair of primers and a probe respectively, for
CC the amplification and detection of a human papilloma virus E6 gene
CC restriction fragment. They are used to exemplify a new method for
CC conducting nucleic acid amplification reactions in the presence of an
CC ionic detergent such as SDS or LLS. The method can be used for enhancing
CC the activity of enzymes such as DNA-directed DNA polymerase, RNA-directed
CC DNA polymerase, DNA-directed RNA polymerase, RNA-hydrolysing enzyme,
CC restriction endonucleases or proteases. (Updated on 27-AUG-2003 to
CC correct OS field.)

SQ Sequence 28 BP; 11 A; 6 C; 6 G; 5 T; 0 U; 0 Other;
Query Match 100.0%; Score 28; DB 2; Length 28;
Best Local Similarity 100.0%; Pred. No. 0.025;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAAACGCAATACAAACACCGTTGTGTG 28
1 GAACAGCAATACAAACACCGTTGTGTG 28
Db

RESULT 2
AAT72300/c
AAT72300 standard; DNA; 28 BP.

AC AAT72300;
09-FEB-1998 (first entry)

DE Human Papillomavirus Type 16 target region.
XX Human Papillomavirus; probe; target region; genital cancer; HPV;
KM cervical smear; ss.
XX Human papillomavirus.
OS BP774518-A2.
XX 21-MAY-1997.
PD 15-NOV-1996; 96BP-00308264.
PF 15-NOV-1995; 95US-0006854P.
PR (GENP-) GEN-PROBE INC.
PA Gordon P, Brentano ST, Carter NM, Hammond PW;
XX WPI; 1997-274349/25.
DR Probes for detection of Human Papillomavirus Type 16 and Type 18 - can
XX distinguish between Type 16 and 18, associated with genital cancers.

PT Claim 1; Page 34; 70pp; English.
PS Novel hybridisation assay probes have been developed comprising an
CC oligonucleotide which will hybridise under selected conditions to Human
CC Papillomavirus (HPV) Type 16 and/or 18 (but not Types 6, 11, 31, 33, 35,
CC 39, 45, 51, 52, or 58) target nucleic acids to form detectable target;
CC probe duplex. The present sequence represents a specifically claimed
CC target region. Oligonucleotides are useful to detect HPV Type 16 and/or
CC 18 in samples e.g. cervical smears, body fluid, and distinguish these
CC from other HPV variants. Papillomaviruses are small DNA viruses and HPV
CC Types 16 and 18 are associated with genital cancers. HPV PCR primers can
CC amplify HPV Type 16 and/or 18 nucleic acid in a sample. HPV Type 16
CC and/or 18 can be detected by adding a probe and detecting probe:target
CC duplex formation; target nucleic acid is optionally amplified. Type 16 or
CC 18 can be specifically detected by amplifying nucleic acids with at least
CC one specifically claimed PCR primer. For Type 18 detection, a helper
CC probe may be used. Oligonucleotides can distinguish target sequences from
CC phylogenetically close HPV Types 16 and 18 whilst available detection

CC sera react with antigens of all papillomaviruses
 XX Sequence 28 BP; 5 A; 6 C; 6 G; 11 T; 0 U; 0 Other;
 SQ

Query Match 100.0%; Score 28; DB 2; Length 28;
 Best Local Similarity 100.0%; Pred. No. 0.025;
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAACAGCAATACAAACCGTGTGTG 28
 DB 28 GAACAGCAATACAAACCGTGTGTG 1

RESULT 3
 AAT72335/C
 ID AAT72335 standard; RNA; 28 BP.
 XX
 AC AAT72335;
 XX
 DT 09-FEB-1998 (first entry)
 XX
 DE Human Papillomavirus Type 16 target region.
 XX
 KM Human Papillomavirus; probe; target region; genital cancer; HPV;
 KM cervical smear; ss.
 XX
 OS Human papillomavirus.
 XX
 PN EP774518-A2.
 XX
 PD 21-MAY-1997.
 XX
 PF 15-NOV-1996; 96EP-00308264.
 XX
 PR 15-NOV-1995; 95US-0006854P.
 XX
 PA (GENP-) GEN-PROBE INC.
 XX
 PI Gordon P, Brentano ST, Carter NM, Hammond PW;
 XX
 DR WPI; 1997-274349/25.
 XX
 PT Probes for detection of Human Papillomavirus Type 16 and Type 18 - can
 PT distinguish between Type 16 and 18, associated with genital cancers.
 PS Claim 1; Page 34; 70pp; English.
 XX
 CC Novel hybridisation assay probes have been developed comprising an
 CC oligonucleotide which will hybridise under selected conditions to Human
 CC Papillomavirus (HPV) Type 16 and/or 18 (but not Types 6, 11, 31, 33, 35,
 CC 39, 45, 51, 52, or 58) target nucleic acids to form detectable target;
 CC probe duplex. The present sequence represents a specifically claimed
 CC target region. Oligonucleotides are useful to detect HPV Type 16 and/or
 CC 18 in samples e.g. cervical smears, body fluid, and distinguish these
 CC from other HPV variants. Papillomaviruses are small DNA viruses and HPV
 CC Types 16 and 18 are associated with genital cancers. HPV PCR primers can
 CC amplify HPV Type 16 and/or 18 nucleic acid in a sample. HPV Type 16
 CC and/or 18 can be detected by adding a probe and detecting probe:target
 CC duplex formation; target nucleic acid is optionally amplified. Type 16 or
 CC 18 can be specifically detected by amplifying nucleic acids with at least
 CC one specifically claimed PCR primer. For Type 18 detection, a helper
 CC probe may be used. Oligonucleotides can distinguish target sequences from
 CC phylogenetically close HPV Types 16 and 18 whilst available detection
 CC sera react with antigens of all papillomaviruses
 CC
 SQ Sequence 28 BP; 5 A; 6 C; 6 G; 0 T; 11 U; 0 Other;
 XX

Query Match 100.0%; Score 28; DB 2; Length 28;
 Best Local Similarity 100.0%; Pred. No. 0.025;
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAACAGCAATACAAACCGTGTGTG 28
 DB 28 GAACAGCAATACAAACCGTGTGTG 1

DB 28 GAACAGCAATACAAACCGTGTGTG 1

RESULT 4
 AAT72299
 ID AAT72299 standard; DNA; 28 BP.
 XX
 AC AAT72299;
 XX
 DT 09-FEB-1998 (first entry)
 XX
 DE Human Papillomavirus Type 16 target region.
 XX
 KM Human Papillomavirus; probe; target region; genital cancer; HPV;
 KM cervical smear; ss.
 XX
 OS Human papillomavirus.
 XX
 PN EP774518-A2.
 XX
 PD 21-MAY-1997.
 XX
 PF 15-NOV-1996; 96EP-00308264.
 XX
 PR 15-NOV-1995; 95US-0006854P.
 XX
 PA (GENP-) GEN-PROBE INC.
 XX
 PI Gordon P, Brentano ST, Carter NM, Hammond PW;
 XX
 DR WPI; 1997-274349/25.
 XX
 PT Probes for detection of Human Papillomavirus Type 16 and Type 18 - can
 PT distinguish between Type 16 and 18, associated with genital cancers.
 XX
 PS Claim 1; Page 34; 70pp; English.
 XX
 CC Novel hybridisation assay probes have been developed comprising an
 CC oligonucleotide which will hybridise under selected conditions to Human
 CC Papillomavirus (HPV) Type 16 and/or 18 (but not Types 6, 11, 31, 33, 35,
 CC 39, 45, 51, 52, or 58) target nucleic acids to form detectable target;
 CC probe duplex. The present sequence represents a specifically claimed
 CC target region. Oligonucleotides are useful to detect HPV Type 16 and/or
 CC 18 in samples e.g. cervical smears, body fluid, and distinguish these
 CC from other HPV variants. Papillomaviruses are small DNA viruses and HPV
 CC Types 16 and 18 are associated with genital cancers. HPV PCR primers can
 CC amplify HPV Type 16 and/or 18 nucleic acid in a sample. HPV Type 16
 CC and/or 18 can be detected by adding a probe and detecting probe:target
 CC duplex formation; target nucleic acid is optionally amplified. Type 16 or
 CC 18 can be specifically detected by amplifying nucleic acids with at least
 CC one specifically claimed PCR primer. For Type 18 detection, a helper
 CC probe may be used. Oligonucleotides can distinguish target sequences from
 CC phylogenetically close HPV Types 16 and 18 whilst available detection
 CC sera react with antigens of all papillomaviruses
 CC
 SQ Sequence 28 BP; 11 A; 6 C; 6 G; 5 T; 0 U; 0 Other;
 XX

Query Match 100.0%; Score 28; DB 2; Length 28;
 Best Local Similarity 100.0%; Pred. No. 0.025;
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAACAGCAATACAAACCGTGTGTG 28
 DB 1 GAACAGCAATACAAACCGTGTGTG 28

RESULT 5
 AAT72334
 ID AAT72334 standard; RNA; 28 BP.
 XX
 AC AAT72334;
 XX
 DT 09-FEB-1998 (first entry)

```

XX DE Human Papillomavirus Type 16 target region.
XX XX
XX KM Human Papillomavirus; probe; target region; genital cancer; HPV;
XX KM cervical smear; ss.
XX OS Human papillomavirus.
XX PN EP74518-A2.
XX PD 21-MAY-1997.
XX PF 15-NOV-1996; 96EP-00308264.
XX PR 15-NOV-1995; 95US-0006854P.
XX PA (GENP-) GEN-PROBE INC.
XX PI Gordon P, Brentano ST, Carter NM, Hammond PW,
XX WPI; 1997-274349/25.
XX DR
XX PT Probes for detection of Human Papillomavirus Type 16 and Type 18 - can
XX PT distinguish between Type 16 and 18, associated with genital cancers.
XX PS Claim 1; Page 34; 70pp; English.
XX CC Novel hybridisation assay probes have been developed comprising an
XX CC oligonucleotide which will hybridise under selected conditions to Human
XX CC Papillomavirus (HPV) Type 16 and/or 18 (but not Types 6, 11, 31, 33, 35,
XX CC 39, 45, 51, 52, or 58) target nucleic acids to form detectable target;
XX CC probe duplex. The present sequence represents a specifically claimed
XX CC target region. Oligonucleotides are useful to detect HPV Type 16 and/or
XX CC 18 in samples e.g. cervical smears, body fluid, and distinguish these
XX CC from other HPV variants. Papillomaviruses are small DNA viruses and HPV
XX CC Types 16 and 18 are associated with genital cancers. HPV PCR primers can
XX CC amplify HPV Type 16 and/or 18 nucleic acid in a sample. HPV Type 16
XX CC and/or 18 can be detected by adding a probe and detecting probe:target
XX CC duplex formation; target nucleic acid is optionally amplified. Type 16 or
XX CC 18 can be specifically detected by amplifying nucleic acids with at least
XX CC one specifically claimed PCR primer. For Type 18 detection, a helper
XX CC probe may be used. Oligonucleotides can distinguish target sequences from
XX CC phylogenetically close HPV Types 16 and 18 whilst available detection
XX CC sera react with antigens of all papillomaviruses
XX SQ Sequence 28 BP; 11 A; 6 C; 6 G; 0 T; 5 U; 0 Other;
XX
Query Match 100.0%; Score 28; DB 2; Length 28;
Best Local Similarity 82.1%; Pred. No. 0.025;
Matches 23; Conservative 5; Mismatches 0; Indels 0; Gaps 0;
QY 1 GAACAGCAATACAAACCGTTGTGTG 28
1 GAACAGCAATACAAACCGTTGTGTG 28
DB 1 GAACAGCAATACAAACCGTTGTGTG 28
RESULT 6
ADX26909
ID ADX26909 standard; DNA; 452 BP.
AC ADX26909;
XX
XX 05-MAY-2005 (first entry)
XX DE Human papillomavirus type 16 B6 delta151 mutant DNA.
XX XX
XX KM cell culture; immortalization; Gynecology and obstetrics; Andrology;
XX KM cancer; neoplasm; mutant; ds.
XX OS Human papillomavirus type 16.
XX OS Synthetic.
XX FH Key Location/Qualifiers

```

```

FT misc_difference 3..4
FT /*tag= a
FT /note= "Wild-type bases 4-25 have been removed"
FT misc_difference 449..450
FT /*tag= b
FT /note= "Wild-type bases 472-474 have been removed"
XX JP2005046117-A.
XX PD 24-FEB-2005.
XX PF 31-JUL-2003; 2003JP-00283911.
XX PR 31-JUL-2003; 2003JP-00283911.
XX PA (UYKA-) UNIT KANAZAWA TLO YG.
XX PA (KOKU-) KOKURITSU GAN CENT SOCHO.
XX DR WPI; 2005-175815/19.
XX PT Immortalized endometrial glandular epithelial cell strain containing
XX PT exogenous immortalizing gene and maintains properties of non-cancer cell
XX PT and typical glandular epithelial cells, useful in field of research of
XX PT reproduction.
XX PS Claim 5; SEQ ID NO 4; 22pp; Japanese.
XX CC The invention relates to a novel immortalized endometrial glandular
XX CC epithelial cell strain in which an exogenous immortalizing gene is
XX CC introduced and the properties of the cell line as non-cancer typical
XX CC glandular epithelial cells, are maintained. The cell line of the
XX CC invention may be useful for research of reproduction, particularly for
XX CC analyzing the adhesion of the fertilized egg during implantation,
XX CC analyzing the effect of steroid hormones, analyzing the mechanism of
XX CC multistep canceration of the inner glandular epithelial membrane by
XX CC genetic engineering of the cell line which is otherwise non-cancerous,
XX CC and analyzing the effect of carcinogenic substances on the cell line. The
XX CC cell line maintains the properties of normal somatic cells and the
XX CC properties of endometrial glandular epithelial cells. The current
XX CC sequence is that of the Human papillomavirus type 16 B6 delta151 mutant
XX CC DNA of the invention.
XX SQ Sequence 452 BP; 156 A; 70 C; 98 G; 128 T; 0 U; 0 Other;
XX
Query Match 100.0%; Score 28; DB 14; Length 452;
Best Local Similarity 100.0%; Pred. No. 0.034;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GAACAGCAATACAAACCGTTGTGTG 28
264 GAACAGCAATACAAACCGTTGTGTG 291
DB 1 GAACAGCAATACAAACCGTTGTGTG 291
RESULT 7
ABQ76223
ID ABQ76223 standard; DNA; 456 BP.
AC ABQ76223;
XX
XX 21-OCT-2002 (first entry)
XX DE Human papillomavirus B6 tumour antigen derived DNA.
XX XX
XX KM Tumour antigen; human; vaccine; cellular immune response; immunogen; B6;
XX KM cancer; tumour; ds.
XX OS Human papillomavirus.
XX OS US6287569-B1.
XX PN 11-SEP-2001.
XX PD 06-APR-1998; 98US-00056105.
XX PF

```


XX 10-APR-1997; 97US-0043467P.
XX (REGC) UNIV CALIFORNIA.
XX K1pps TJ, Wu Y;
XX WPI; 1998-583198/49.
XX
XX
XX Generating cellular immune response in patient to target protein -
PT comprises introducing vector with nucleotide sequence encoding immunogen
PT comprising protein processing signal into cell of patient.
XX
XX Disclosure; Col 91-92; 61pp; English.
XX
XX This invention describes a novel method for generating a cellular immune
CC response in a patient to a target protein or its fragment. The method
CC involves introducing a vector containing a nucleotide sequence encoding a
CC chimeric immunogen comprising a protein processing signal and the target
CC protein or its fragment. The immunogen is produced by the cells and
CC processed so that the target protein or its fragment is presented to the
CC patient's immune system and a cellular immune response is initiated. The
CC method and vectors can be used as a form of vaccination and could be used
CC to generate a cellular immune response in patients to, e.g. cancerous
CC tumours. The cellular immune response is the predominant immune response
CC in the patient. This sequence represents a DNA fragment which encodes a
CC tumour antigen from human papillomavirus B6 described in the method of
CC the invention. Note: The information in this spec has been previously
CC disclosed in WO19945444 however this spec contained no sequence
XX
XX
XX Sequence 456 BP; 156 A; 70 C; 101 G; 129 T; 0 U; 0 Other;
SQ
Query Match 100.0%; Score 28; DB 2; Length 456;
Best Local Similarity 100.0%; Pred. No. 0.034;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GAACGAGATACACAAACCGTGTGTG 28
DB 265 GAACGAGATACACAAACCGTGTGTG 292
RESULT 8
ID ADC37149 standard; DNA; 459 BP.
XX
XX ADC37149;
AC
XX
XX 18-DEC-2003 (first entry)
DT
XX
XX 459bp HPV16 subgenomic fragment DNA sequence.
DE
XX
XX immortalised non-tumorigenic; human Schwann; schwannoma cell line;
KM screening cancer chemotherapeutic; antineoplastic activity;
KM carcinogenicity; neuroprotective; neurodegeneration; SV40;
KM Simian Virus 40; adenovirus; human papilloma virus; B6; E7; cytostatic;
KM gene therapy; cancer; HPV16; ds.
XX
XX Human papillomavirus type 16.
OS
XX
XX WO2003073996-A2.
PN
XX
XX 12-SEP-2003.
PD
XX
XX 03-MAR-2003; 2003WO-US006314.
PF
XX
XX 01-MAR-2002; 2002US-0361528P.
PR
XX
XX (HOU-S-) HOUSE EAR INST.
PA
XX
XX Hung G, Li X;
PI
XX
XX WPI; 2003-722006/68.

XX
XX Producing an immortalized non-tumorigenic human Schwann or schwannoma
PT cell line for treating cancer or neurodegenerative disorder comprises
PT selecting for immortalized cells that express the exogenous immortalizing
PT gene.
XX
XX Disclosure; SEQ ID NO 1; 29pp; English.
XX
XX The invention relates to a the production of an immortalised non-
CC tumorigenic human Schwann or schwannoma cell line comprising: providing
CC a cell culture of human Schwann or schwannoma cells, introducing a
CC polynucleotide comprising an exogenous immortalising gene into the cells,
CC and selecting for immortalised cells that express the exogenous
CC immortalising gene and retain phenotypic properties of Schwann or
CC schwannoma cells. The invention further comprises: determining the effect
CC of a pharmacological agent on human Schwann or schwannoma cells;
CC screening cancer chemotherapeutic and antineoplastic activity of an agent
CC ; testing carcinogenicity of an agent; screening the neuroprotective
CC activity of an agent; treating neurodegeneration in a patient; and a kit
CC for screening a pharmacological agent on schwannoma cells. In the method
CC of producing an immortalised non-tumorigenic human Schwann or schwannoma
CC cell line, the polynucleotide is a subgenomic fragment of a virus
CC consisting of SV40 (Simian Virus 40), adenovirus and human papilloma
CC virus, comprising the B6 and E7 genes of the human papilloma virus
CC consisting of types 18, 31, 33 or 35, but preferably 16. The immortalised
CC non-tumorigenic human Schwann or schwannoma cell line has cytostatic and
CC neuroprotective activity and be useful in treating disorders by gene
CC therapy. The method is useful for producing an immortalised non-
CC tumorigenic human Schwann or schwannoma cell line useful for screening
CC cancer chemotherapeutic and antineoplastic activity of an agent, testing
CC carcinogenicity of an agent and screening a neuroprotective activity of
CC an agent and for preparing a composition for diagnosing or treating
CC neurodegenerative disorders and cancer. This polynucleotide represents
CC the DNA of a 459bp HPV16 sequence, a subgenomic fragment used in the
CC production of the immortalised non-tumorigenic human Schwann or
CC schwannoma cell line of the invention.
XX
XX
XX Sequence 459 BP; 157 A; 71 C; 100 G; 131 T; 0 U; 0 Other;
SQ
Query Match 100.0%; Score 28; DB 10; Length 459;
Best Local Similarity 100.0%; Pred. No. 0.034;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GAACGAGATACACAAACCGTGTGTG 28
DB 266 GAACGAGATACACAAACCGTGTGTG 293
RESULT 9
ID ADF09607 standard; CDNA; 477 BP.
XX
XX ADF09607;
AC
XX
XX 12-FEB-2004 (first entry)
DT
XX
XX Human papillomavirus 16 B6 cDNA SEQ ID NO:108.
DE
XX
XX human; protein-protein interaction; virucide; cytostatic; vaccine;
KM human papilloma virus; HPV; cancer; ss; gene.
KM
XX
XX Human papillomavirus.
OS
XX
XX WO2003068940-A2.
FN
XX
XX 21-AUG-2003.
PD
XX
XX 14-FEB-2003; 2003WO-US004594.
PF
XX
XX 14-FEB-2002; 2002US-0356911P.
PR
XX
XX (CURA-) CURAGEN CORP.
PA
XX
XX (HOF) HOFFMANN LA ROCHE INC.

CC silencing in a mammalian cell of the expression of an exogenous gene of
 CC viral origin. The method comprises introducing into the cell a small
 CC interfering RNA (siRNA) construct that is homologous to a part of the
 CC mRNA sequence of the gene. The method is useful for the selective post-
 CC transcriptional silencing of an exogenous gene of viral origin (e.g. an
 CC oncogene or human papilloma virus (HPV) gene) in a mammalian cell. The
 CC method or the siRNA is particularly useful for treating cancer, human
 CC cervical cancer, human immunodeficiency virus (HIV), smallpox, flu,
 CC common cold, or a disease caused by a HPV (e.g. genital warts, cervical
 CC cancer, penile cancer, malignant squamous cell carcinomas or verruca
 CC vulgaris). An siRNA construct or vector is useful for use as a medicament
 CC for the diseases mentioned. The polynucleotide sequence of the invention
 CC can be used to treat disorder by gene therapy. This polynucleotide
 CC sequence represents the DNA of a HPV siRNA sequence of the invention
 CC
 SQ Sequence 477 BP, 167 A, 75 C, 104 G, 131 T, 0 U, 0 Other;
 Query Match 100.0%; Score 28; DB 10; Length 477;
 Best Local Similarity 100.0%; Pred. No. 0.034;
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GAACAGCAATACAAACCGTTGTG 28
 Db 286 GAACAGCAATACAAACCGTTGTG 313
 RESULT 12
 ADX26908
 ID ADX26908 standard; DNA; 477 BP.
 AC ADX26908;
 XX
 DT 05-MAY-2005 (first entry)
 DE Human papillomavirus type 16 E6 wild-type gene.
 XX
 KM cell culture; immortalization; gynecology and obstetrics; Andrology;
 KM cancer; neoplasms; ds.
 KW
 OS Human papillomavirus type 16.
 XX
 PN JP2005046117-A.
 XX
 PD 24-FEB-2005.
 XX
 PF 31-JUL-2003; 2003JP-00283911.
 XX
 PR 31-JUL-2003; 2003JP-00283911.
 XX
 PA (UYKA-) UNIV KANAZAMA TLO YG.
 PA (KOKU-) KOKURITSU GAN CENT SOCHO.
 XX
 DR WPI; 2005-175815/19.
 XX
 XX
 PT Immortalized endometrial glandular epithelial cell strain containing
 PT exogenous immortalizing gene and maintains properties of non-cancer cell
 PT and typical glandular epithelial cells, useful in field of research of
 PT reproduction.
 XX
 PS Claim 4; SEQ ID NO 3; 22pp; Japanese.
 XX
 CC The invention relates to a novel immortalized endometrial glandular
 CC epithelial cell strain in which an exogenous immortalizing gene is
 CC introduced and the properties of the cell line as non-cancer typical
 CC glandular epithelial cells, are maintained. The cell line of the
 CC invention may be useful for research of reproduction, particularly for
 CC analyzing the adhesion of the fertilized egg during implantation,
 CC analyzing the effect of steroid hormones, analyzing the mechanism of
 CC multistep canceration of the inner glandular epithelial membrane by
 CC genetic engineering of the cell line which is otherwise non-cancerous,
 CC and analyzing the effect of carcinogenic substances on the cell line. The
 CC cell line maintains the properties of normal somatic cells and the
 CC properties of endometrial glandular epithelial cells. The current

CC sequence is that of the Human papillomavirus type 16 E6 wild-type gene of
 CC the invention.
 XX
 SQ Sequence 477 BP, 167 A, 76 C, 103 G, 131 T, 0 U, 0 Other;
 Query Match 100.0%; Score 28; DB 14; Length 477;
 Best Local Similarity 100.0%; Pred. No. 0.034;
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GAACAGCAATACAAACCGTTGTG 28
 Db 286 GAACAGCAATACAAACCGTTGTG 313
 RESULT 13
 AEA5113
 ID AEA5113 standard; DNA; 477 BP.
 AC AEA5113;
 XX
 DT 11-AUG-2005 (first entry)
 DE Human papillomavirus type 16 E6/E7 gene fragment, SEQ ID NO: 17.
 XX
 KM Delivery mechanism; gene therapy; cytostatic; vulnery; virocidic;
 KM injury; infection; cancer; gastrointestinal disease;
 KM gynecology and obstetrics; tumor; colorectal tumor;
 KM uterine cervix tumor; squamous cell carcinoma; neoplasms; E7 gene; ds;
 KM E6 gene.
 XX
 OS Human papillomavirus.
 XX
 PN WO2005051431-A1.
 XX
 PD 09-JUN-2005.
 XX
 PF 25-NOV-2004; 2004WO-GB004979.
 XX
 PR 25-NOV-2003; 2003GB-00027409.
 PR 05-MAR-2004; 2004US-0549919P.
 XX
 PA (MILN/) MILNER A J.
 PA
 PI Milner AJ;
 XX
 DR WPI; 2005-405310/41.
 XX
 XX
 PT Composition useful for delivering an agent into a cell comprises the
 PT agent, a transfer agent and a solid or colloidal carrier medium.
 XX
 PS Disclosure; SEQ ID NO 17, 56pp; English.
 XX
 CC The present invention relates to a method and composition comprising a
 CC transfer agent and a solid or colloidal carrier medium for delivering
 CC biological agents into cells. The invention is useful for the treatment
 CC of cancer, tumor, carcinoma of cutaneous, squamous or cervical epithelia
 CC and colorectal carcinoma, wounds, burns and scars. The invention is also
 CC useful in gene therapy. The present sequence is the human papillomavirus
 CC type 16 E6/E7 gene fragment. This sequence is useful in a method for
 CC treating carcinomas.
 CC
 SQ Sequence 477 BP, 167 A, 75 C, 104 G, 131 T, 0 U, 0 Other;
 Query Match 100.0%; Score 28; DB 14; Length 477;
 Best Local Similarity 100.0%; Pred. No. 0.034;
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GAACAGCAATACAAACCGTTGTG 28
 Db 286 GAACAGCAATACAAACCGTTGTG 313
 RESULT 14

AAT31834
 ID AAT31834 standard; DNA; 519 BP.
 XX
 AC AAT31834;
 XX
 DT 27-AUG-2003 (revised)
 DT 11-JAN-1997 (first entry)
 XX
 DE Human papilloma virus E6/E7 protein variant.
 XX
 KM Human papilloma virus; E6, E7; deletion mutant; HPV; immune response;
 KM humoral immune response; cellular immune response; vaccine; ss.
 XX
 OS Human papillomavirus.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..519
 FT /*tag= a
 FT /product= "E6/E7 fusion protein (deletion mutant)."
 FT 109..111
 FT /*tag= b
 FT /transl_except= TAC encodes Ile
 FT 115..117
 FT /*tag= c
 FT /transl_except= AAA encodes Cys
 FT 118..120
 FT /*tag= d
 FT /transl_except= CCG encodes Gln
 FT 121..123
 FT /*tag= e
 FT /transl_except= TTG encodes Lys
 FT 124..126
 FT /*tag= f
 FT /transl_except= TGT encodes Pro
 FT 127..129
 FT /*tag= g
 FT /transl_except= GAT encodes Tyr
 FT 130..132
 FT /*tag= h
 FT /transl_except= TTG encodes Asn
 FT 133..135
 FT /*tag= i
 FT /transl_except= TTA encodes Lys
 FT 136..138
 FT /*tag= j
 FT /transl_except= ATT encodes Pro
 FT 139..141
 FT /*tag= k
 FT /transl_except= AGG encodes Leu
 FT 142..144
 FT /*tag= l
 FT /transl_except= TGT encodes Cys
 FT 145..147
 FT /*tag= m
 FT /transl_except= ATT encodes Asp
 FT 148..150
 FT /*tag= n
 FT /transl_except= AAC encodes Leu
 FT 151..153
 FT /*tag= o
 FT /transl_except= TGT encodes Leu
 FT 154..156
 FT /*tag= p
 FT /transl_except= CAA encodes Ile
 FT 157..159
 FT /*tag= q
 FT /transl_except= AAG encodes Arg
 FT 160..162
 FT /*tag= r
 FT /transl_except= CCA encodes Cys
 XX
 PN WO9619496-A1.
 XX

PD 27-JUN-1996.
 XX
 PF 20-DEC-1995; 95MO-AU000068.
 XX
 PR 20-DEC-1994; 94AU-00000157.
 XX
 PA (CSLC-) CSL LTD.
 PA (UYOU) UNIV QUEBENSIAIND.
 XX
 PI Edwards SJ, Cox J, Webb EA, Frazer I;
 XX
 DR WPI, 1996-309518/31.
 DR P-PSDB; AAR97562.
 XX
 PT Vaccine variants of human papilloma virus antigens - contain variants of
 PT E6 and/or E7 protein, pref. deletion mutants, and are used to treat or
 PT prevent HPV infection.
 XX
 PS Example 3; Page 17; 37pp; English.
 XX
 CC A variant of the human papilloma virus (HPV) E6 or E7 protein which
 CC elicits a humoral and/or cellular immune response against HPV can be used
 CC in vaccines against HPV or to treat HPV infection. The variant is
 CC preferably a deletion mutant comprising at least half, and preferably two
 CC thirds of full length E6 or E7 protein starting from the N- or C-
 CC terminal, or is a full length E6 moiety fused to a full length E7 moiety.
 CC The variant optionally has a linkage moiety and a foreign protein or
 CC peptide which facilitates the purification of, and enhances the
 CC immunogenicity of, the fusion protein. This sequence encodes a fusion
 CC between the C-terminal end of E6 and the N-terminal end of E7. The
 CC protein is also a deletion mutant generated from the sequence described
 CC in AAT31833. (Updated on 27-AUG-2003 to correct OS field.)
 XX
 SQ Sequence 519 BP; 175 A; 92 C; 113 G; 139 T; 0 U; 0 Other;
 QY 1 GAACAGCAATACAAACCGTTGTG 28
 Db 100 GAACAGCAATACAAACCGTTGTG 127
 RESULT 15
 ID ADF31984 standard; DNA; 543 BP.
 XX
 AC ADF31984;
 XX
 DT 12-FEB-2004 (first entry)
 DT 12-FEB-2004 (first entry)
 XX
 DE Human papillomavirus fusion gene encoding sequence.
 XX
 KM human papillomavirus; cervix cancer; ds.
 KM
 OS Homo sapiens.
 OS
 FH Key Location/Qualifiers
 FT CDS 1..543
 FT /*tag= a
 XX
 PN CN1361583-A.
 XX
 PD 27-NOV-2002.
 XX
 PF 24-APR-2002; 2002CN-00117143.
 XX
 PR 24-APR-2002; 2002CN-00117143.
 XX
 PA (ONCO-) INST ONCOLOGY TUMOR HOSPITAL CHINESE ACA.
 XX
 PI Zhao Q;
 XX

XX WPI; 2003-258260/26.
 DR P-PSDB; ADF31985.
 XX Human papillomavirus E6/E7 fusion gene and its efficient expression
 PT carrier and fusion protein vaccine.
 XX
 PS Claim 3; SEQ ID NO 1; 16pp; Chinese.
 CC The present invention relates to human papillomavirus E6/E7 fusion gene,
 CC its preparing process, the process for configuring the efficient
 CC expression carrier containing the gene and resultant expression carrier,
 CC the fusion protein prepared from the gene, and the application of the
 CC fusion gene and expression protein to medical science and medicine to
 CC treat cervix cancer are disclosed. The present sequence represents the
 CC human papillomavirus fusion gene encoding sequence.
 CC
 SQ Sequence 543 BP; 185 A; 89 C; 118 G; 151 T; 0 U; 0 Other;
 Query Match 100.0%; Score 28; DB 10; Length 543;
 Best Local Similarity 100.0%; Pred. No. 0.035;
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GAACAGCAATACAAACCGTTGTG 28
 Db 286 GAACAGCAATACAAACCGTTGTG 313
 RESULT 16
 AAQ75470
 ID AAQ75470 standard; DNA; 570 BP.
 AC AAQ75470;
 XX 16-OCT-2003 (revised)
 DT 25-MAR-2003 (revised)
 DT 28-JUN-1995 (first entry)
 XX
 XX HPV16 E6/E7 encoding region.
 DE
 XX HPV; HPV16; E6 protein; E7 protein; diagnosis; cervical dysplasia;
 KM cervix cancer; de.
 XX
 OS Human papillomavirus; strain 16.
 XX
 XX Key Location/Qualifiers
 FH CDS 2..478
 FT /*tag= a
 FT /label= B6_encoding_region
 FT 481..570
 FT /*tag= b
 FT /label= E7_encoding_region
 PN
 XX WO9426934-A2.
 PD 24-NOV-1994.
 XX
 XX 06-MAY-1994; 94WO-US005085.
 PF
 XX 06-MAY-1993; 93US-00058920.
 PR
 PA (BAXT) BAXTER DIAGNOSTICS INC.
 PI Brown JT;
 XX WPI; 1995-006821/01.
 DR P-PSDB; AAR63865.
 XX
 XX Human papilloma virus detection assay - by amplification using self
 PT sustained sequence replication and hybridisation with a detector probe.
 XX
 PS Disclosure; Page 24-26; 79pp; English.
 XX

CC The sequences of the E6 and E7 polypeptide-encoding regions of human
 CC papillomavirus (HPV) 16 and 18 are given in AAQ75470-71 and the encoded
 CC proteins in AAR63865-66, respectively. Probes and primers based on these
 CC sequences were used for HPV infection diagnosis; expression of E6 and E7
 CC is diagnostic for cervical cancer or pre-malignant states. (Updated on
 CC 25-MAR-2003 to correct PN field.) (Updated on 16-Oct-2003 to standardise
 CC OS field)
 CC
 SQ Sequence 570 BP; 199 A; 93 C; 119 G; 159 T; 0 U; 0 Other;
 Query Match 100.0%; Score 28; DB 2; Length 570;
 Best Local Similarity 100.0%; Pred. No. 0.035;
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GAACAGCAATACAAACCGTTGTG 28
 Db 287 GAACAGCAATACAAACCGTTGTG 314
 RESULT 17
 AAN91600
 ID AAN91600 standard; DNA; 712 BP.
 AC AAN91600;
 XX 27-AUG-2003 (revised)
 DT 17-JUL-1990 (first entry)
 XX
 XX Partial nucleotide sequence (5' end) of human papilloma virus (HPV) type
 DE 16 (HPV-16).
 XX
 XX Human papilloma virus; type 16; in situ hybridisation assay;
 KM cellular smear; benign cervical wart; cervical cancer.
 XX
 XX Human papillomavirus.
 OS
 XX WO8902934-A.
 PN
 XX 06-APR-1989.
 PD
 XX 30-SEP-1988; 88WO-US003367.
 PF
 XX 02-OCT-1987; 87US-00103979.
 PR
 PA (MICR-) MICROPROBE CORP.
 PI Schwartz DE, Adams TH;
 XX WPI; 1989-114406/15.
 DR
 XX Hybridisation test for human papilloma virus in cell smears - by reaction
 PT with long labelled probe specific for particular virus types, esp. for
 PT examining cervical smears.
 XX
 PS Disclosure; Page 7; 39pp; English.
 XX
 XX The patent is for a rapid in situ hybridisation assay for detecting and
 CC typing human papilloma virus (HPV) in non-frozen cellular smears fixed to
 CC a support in absence of aldehyde-based crosslinking reagents. The assay
 CC comprises: (1) combining nucleic acid in the sample with at least one
 CC detectable probe able to hybridise with 1 or more HPV types; and (2)
 CC detecting presence or absence of hybrid complexes. Opt. several probes
 CC are used, eg one for HPV types 6 and 11, associated with benign warts,
 CC and one for types 16, 18, 31, 33 and 35, associated with cervical cancer.
 CC The assay can differentiate between HPV types. It is esp. used as a
 CC secondary test. The probes can be synthesised or cloned. (Updated on 27-
 CC AUG-2003 to correct OS field.)
 CC
 SQ Sequence 712 BP; 253 A; 120 C; 155 G; 184 T; 0 U; 0 Other;
 Query Match 100.0%; Score 28; DB 1; Length 712;
 Best Local Similarity 100.0%; Pred. No. 0.036;
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAACAGCAATACAAACCGTTGTG 28
DB 368 GAACAGCAATACAAACCGTTGTG 395

RESULT 18

ADO44067
ID ADO44067 standard; DNA; 747 BP.

XX ADO44067;

DT 15-JUL-2004 (first entry)

DE Nucleotide sequence of an E7B6 fusion protein.

XX E6 protein; E7 protein; fusion protein; HPV16; HPV-associated cancer;
KM cervical cancer; immune response; lower gastrointestinal tract cancer;
KM anal cancer; reproductive system cancer; penile cancer; vulvar cancer;
KM gene; ss.

XX Human papillomavirus type 16.

OS Synthetic.

FH Key Location/Qualifiers

FT CDS 1..747

FT /tag= a

FT /product= "E7B6 fusion protein"

PN MO2004030636-A2.

PD 15-APR-2004.

PF 02-OCT-2003; 2003MO-US031726.

PR 03-OCT-2002; 2002US-0415929P.

PA (AMHP) WYETH HOLDINGS CORP.

PI Smith L, Casasetti MC;

DR WPI; 2004-316328/29.

DR P-PSDB; ADO44066.

XX New polypeptide comprising human papillomavirus E6 and E7 polypeptides,
PT useful for treating or preventing human papillomavirus (HPV)-associated
PT cancers, e.g. cervical cancer.

XX Example 1; Page 73; 101pp; English.

XX The present sequence encodes an E7B6 fusion protein, comprising wild type
CC E7 and B6 polypeptides from human papillomavirus type 16 (HPV16). The
CC specification describes human papillomavirus B6 and E7 polypeptides,
CC where the E7 polypeptide has mutations at any one or more of the amino
CC acids corresponding to amino acids 24, 26 or 91 of the sequence given in
CC ADO44073 and the B6 polypeptide has no mutations or has mutations at any
CC one or more of the amino acids corresponding to amino acids 63 or 106 of
CC the sequence given in ADO44072. The polypeptides of the invention are
CC useful for treating or preventing human papillomavirus (HPV)-associated
CC cancers, such as cervical cancer. The fusion proteins and nucleic acids
CC encoding the fusion proteins are useful for generating immune responses
CC against HPV. They are also useful for treating lower gastrointestinal
CC tract cancers, e.g. anal cancer, and other cancers of the reproductive
CC system, including penile and vulvar cancer.

XX Sequence 747 BP; 247 A; 132 C; 166 G; 202 T; 0 U; 0 Other;

XX Query Match 100.0%; Score 28; DB 12; Length 747;

XX Best Local Similarity 100.0%; Pred. No. 0.036;

XX Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAACAGCAATACAAACCGTTGTG 28

DB 556 GAACAGCAATACAAACCGTTGTG 583

RESULT 19

ADO44065
ID ADO44065 standard; DNA; 747 BP.

XX ADO44065;

DT 15-JUL-2004 (first entry)

DE Nucleotide sequence of a fusion protein designated B6E7pentM.

XX E6 protein; E7 protein; fusion protein; HPV16; HPV-associated cancer;
KM cervical cancer; immune response; lower gastrointestinal tract cancer;
KM anal cancer; reproductive system cancer; penile cancer; vulvar cancer;
KM gene; ss.

XX Human papillomavirus type 16.

OS Synthetic.

FH Key Location/Qualifiers

FT CDS 1..747

FT /tag= a

FT /product= "B6E7pentM fusion protein"

PN MO2004030636-A2.

PD 15-APR-2004.

PF 02-OCT-2003; 2003MO-US031726.

PR 03-OCT-2002; 2002US-0415929P.

PA (AMHP) WYETH HOLDINGS CORP.

PI Smith L, Casasetti MC;

DR WPI; 2004-316328/29.

DR P-PSDB; ADO44064.

XX New polypeptide comprising human papillomavirus E6 and E7 polypeptides,
PT useful for treating or preventing human papillomavirus (HPV)-associated
PT cancers, e.g. cervical cancer.

XX Claim 24; Page 71; 101pp; English.

XX The present sequence encodes a fusion protein, comprising E6 and E7
CC polypeptides from human papillomavirus type 16 (HPV16). The fusion
CC protein is designated B6E7pentM, and comprises an B6 amino terminus
CC (where residues 63 and 106 have been replaced with glycine) and an E7
CC carboxy terminus (where residues 24, 26 and 91 have been replaced with
CC glycine). B6E7pentM is representative of fusion proteins of the
CC invention. The specification describes human papillomavirus B6 and E7
CC polypeptides, where the E7 polypeptide has mutations at any one or more
CC of the amino acids corresponding to amino acids 24, 26 or 91 of the
CC sequence given in ADO44073 and the B6 polypeptide has no mutations or has
CC mutations at any one or more of the amino acids corresponding to amino
CC acids 63 or 106 of the sequence given in ADO44072. The polypeptides of
CC the invention are useful for treating or preventing human papillomavirus
CC (HPV)-associated cancers, such as cervical cancer. The fusion proteins
CC and nucleic acids encoding the fusion proteins are useful for generating
CC immune responses against HPV. They are also useful for treating lower
CC gastrointestinal tract cancers, e.g. anal cancer, and other cancers of
CC the reproductive system, including penile and vulvar cancer.

XX Sequence 747 BP; 246 A; 132 C; 171 G; 198 T; 0 U; 0 Other;

XX Query Match 100.0%; Score 28; DB 12; Length 747;

XX Best Local Similarity 100.0%; Pred. No. 0.036;

XX Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAACAGCAATACAAACCGTTGTG 28

Db 265 GAACAGCAATATCAACAAACCGTTGTG 292

RESULT 20

ID ADO44061 standard; DNA; 747 BP.

AC ADO44061;

DT 15-JUL-2004 (first entry)

DE Nucleotide sequence of an E6E7 fusion protein.

XX E6 protein; E7 protein; fusion protein; HPV16; HPV-associated cancer;
KM cervical cancer; immune response; lower gastrointestinal tract cancer;
KM anal cancer; reproductive system cancer; penile cancer; vulvar cancer;
KW gene; ss.

XX Human papillomavirus type 16.
OS Synthetic.

XX Key Location/Qualifiers
FT CDS 1..747
FT /*tag= a
FT /product= "E6E7 fusion protein"

PN WO2004030636-A2.

PD 15-APR-2004.

PF 02-OCT-2003; 2003WC-US031726.

PR 03-OCT-2002; 2002US-0415929P.

PA (AMHP) WYETH HOLDINGS CORP.

PI Smith L, Cassetti MC;

DR WPI; 2004-316328/29.

DR P-PSDB; ADO44060.

XX New polypeptide comprising human papillomavirus E6 and E7 polypeptides,
PT useful for treating or preventing human papillomavirus (HPV)-associated
PT cancers, e.g. cervical cancer.

XX Example 1; Page 68; 101pp; English.

XX The present sequence encodes an E6E7 fusion protein, comprising wild type
CC E6 and E7 polypeptides from human papillomavirus type 16 (HPV16). The
CC specification describes human papillomavirus E6 and E7 polypeptides,
CC where the E7 polypeptide has mutations at any one or more of the amino
CC acids corresponding to amino acids 24, 26 or 91 of the sequence given in
CC ADO44073 and the E6 polypeptide has no mutations or has mutations at any
CC one or more of the amino acids corresponding to amino acids 63 or 106 of
CC the sequence given in ADO44072. The polypeptides of the invention are
CC useful for treating or preventing human papillomavirus (HPV)-associated
CC cancers, such as cervical cancer. The fusion proteins and nucleic acids
CC encoding the fusion proteins are useful for generating immune responses
CC against HPV. They are also useful for treating lower gastrointestinal
CC tract cancer, e.g. anal cancer, and other cancers of the reproductive
CC system, including penile and vulvar cancer.

XX Sequence 747 BP; 247 A; 132 C; 166 G; 202 T; 0 U; 0 Other;

Query Match 100.0%; Score 28; DB 12; Length 747;

Best Local Similarity 100.0%; Pred. No. 0.036;

Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAACAGCAATATCAACAAACCGTTGTG 28

Db 265 GAACAGCAATATCAACAAACCGTTGTG 292

RESULT 21

ID ADO44063 standard; DNA; 747 BP.

AC ADO44063;

DT 15-JUL-2004 (first entry)

DE Nucleotide sequence of a fusion protein designated E6E7TetM.

XX E6 protein; E7 protein; fusion protein; HPV16; HPV-associated cancer;
KM cervical cancer; immune response; lower gastrointestinal tract cancer;
KM anal cancer; reproductive system cancer; penile cancer; vulvar cancer;
KW gene; ss.

XX Human papillomavirus type 16.
OS Synthetic.

XX Key Location/Qualifiers
FT CDS 1..747
FT /*tag= a
FT /product= "E6E7TetM fusion protein"

PN WO2004030636-A2.

PD 15-APR-2004.

PF 02-OCT-2003; 2003WC-US031726.

PR 03-OCT-2002; 2002US-0415929P.

PA (AMHP) WYETH HOLDINGS CORP.

PI Smith L, Cassetti MC;

DR WPI; 2004-316328/29.

DR P-PSDB; ADO44062.

XX New polypeptide comprising human papillomavirus E6 and E7 polypeptides,
PT useful for treating or preventing human papillomavirus (HPV)-associated
PT cancers, e.g. cervical cancer.

XX Claim 24; Page 70; 101pp; English.

XX The present sequence encodes a fusion protein, comprising E6 and E7
CC polypeptides from human papillomavirus type 16 (HPV16). The fusion
CC protein is designated E6E7TetM, and comprises an E6 amino terminus (where
CC residues 63 and 106 have been replaced with glycine) and an E7 carboxy
CC terminus (where residues 24 and 26 have been replaced with glycine).
CC E6E7TetM is representative of fusion proteins of the invention. The
CC specification describes human papillomavirus E6 and E7 polypeptides,
CC where the E7 polypeptide has mutations at any one or more of the amino
CC acids corresponding to amino acids 24, 26 or 91 of the sequence given in
CC ADO44073 and the E6 polypeptide has no mutations or has mutations at any
CC one or more of the amino acids corresponding to amino acids 63 or 106 of
CC the sequence given in ADO44072. The polypeptides of the invention are
CC useful for treating or preventing human papillomavirus (HPV)-associated
CC cancers, such as cervical cancer. The fusion proteins and nucleic acids
CC encoding the fusion proteins are useful for generating immune responses
CC against HPV. They are also useful for treating lower gastrointestinal
CC tract cancer, e.g. anal cancer, and other cancers of the reproductive
CC system, including penile and vulvar cancer.

XX Sequence 747 BP; 246 A; 132 C; 170 G; 199 T; 0 U; 0 Other;

Query Match 100.0%; Score 28; DB 12; Length 747;

Best Local Similarity 100.0%; Pred. No. 0.036;

Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAACAGCAATATCAACAAACCGTTGTG 28

Db 265 GAACAGCAATATCAACAAACCGTTGTG 292

```
RESULT 22
ADO44069
ID ADO44069 standard; DNA; 747 BP.
XX
AC ADO44069;
XX
DT 15-JUL-2004 (first entry)
XX
DE Nucleotide sequence of a fusion protein designated E7B6rectm.
XX
KW E6 protein; E7 protein; fusion protein; HPV16; HPV-associated cancer;
KM cervical cancer; immune response; lower gastrointestinal tract cancer;
KM anal cancer; reproductive system cancer; penile cancer; vulvar cancer;
KM gene; ss.
XX
OS Human papillomavirus type 16.
XX
OS Synthetic.
XX
FH Key 1. .747
FT CDS /tag= a
FT /product= "E7B6rectm fusion protein"
XX
XX MO2004030636-A2.
XX
PD 15-APR-2004.
XX
PF 02-OCT-2003; 2003MO-US031726.
XX
PR 03-OCT-2002; 2002US-0415929P.
XX
PA (AMHP ) WYETH HOLDINGS CORP.
XX
PI Smith L, Cassetti MC;
XX
PI WPI; 2004-316328/29.
XX
DR P-PSDB; ADO44068.
XX
XX New polypeptide comprising human papillomavirus E6 and E7 polypeptides,
PT useful for treating or preventing human papillomavirus (HPV)-associated
PT cancers, e.g. cervical cancer.
XX
XX Claim 24; Page 74-75; 101pp; English.
XX
PS The present sequence encodes a fusion protein, comprising E7 and E6
CC polypeptides from human papillomavirus type 16 (HPV16). The fusion
CC protein is designated E7B6rectm, and comprises an E7 amino terminus (where
CC residues 24 and 26 have been replaced with glycine) and an E6 carboxy
CC terminus (where residues 63 and 106 have been replaced with glycine).
CC E7B6rectm is representative of fusion proteins of the invention. The
CC specification describes human papillomavirus E6 and E7 polypeptides,
CC where the E7 polypeptide has mutations at any one or more of the amino
CC acids corresponding to amino acids 24, 26 or 91 of the sequence given in
CC ADO44073 and the E6 polypeptide has no mutations or has mutations at any
CC one or more of the amino acids corresponding to amino acids 63 or 106 of
CC the sequence given in ADO44072. The polypeptides of the invention are
CC useful for treating or preventing human papillomavirus (HPV)-associated
CC cancers, such as cervical cancer. The fusion proteins and nucleic acids
CC encoding the fusion proteins are useful for generating immune responses
CC against HPV. They are also useful for treating lower gastrointestinal
CC tract cancers, e.g. anal cancer, and other cancers of the reproductive
CC system, including penile and vulvar cancer.
XX
XX Sequence 747 BP; 246 A; 132 C; 170 G; 199 T; 0 U; 0 Other;
SQ
Query Match 100.0%; Score 28; DB 12; Length 747;
Best Local Similarity 100.0%; Pred. No. 0.036;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
DB 556 GAACAGCAATACAAACCGTTGTGTG 583
RESULT 23
ADO44071
ID ADO44071 standard; DNA; 747 BP.
XX
AC ADO44071;
XX
DT 15-JUL-2004 (first entry)
XX
DE Nucleotide sequence of a fusion protein designated E7B6Pentm.
XX
KW E6 protein; E7 protein; fusion protein; HPV16; HPV-associated cancer;
KM cervical cancer; immune response; lower gastrointestinal tract cancer;
KM anal cancer; reproductive system cancer; penile cancer; vulvar cancer;
KM gene; ss.
XX
OS Human papillomavirus type 16.
XX
OS Synthetic.
XX
FH Key 1. .747
FT CDS /tag= a
FT /product= "E7B6Pentm fusion protein"
XX
XX MO2004030636-A2.
XX
PD 15-APR-2004.
XX
PF 02-OCT-2003; 2003MO-US031726.
XX
PR 03-OCT-2002; 2002US-0415929P.
XX
PA (AMHP ) WYETH HOLDINGS CORP.
XX
PI Smith L, Cassetti MC;
XX
PI WPI; 2004-316328/29.
XX
DR P-PSDB; ADO44070.
XX
XX New polypeptide comprising human papillomavirus E6 and E7 polypeptides,
PT useful for treating or preventing human papillomavirus (HPV)-associated
PT cancers, e.g. cervical cancer.
XX
XX Claim 24; Page 76; 101pp; English.
XX
PS The present sequence encodes a fusion protein, comprising E7 and E6
CC polypeptides from human papillomavirus type 16 (HPV16). The fusion
CC protein is designated E7B6Pentm, and comprises an E7 amino terminus
CC (where residues 24, 26 and 91 have been replaced with glycine) and an E6
CC carboxy terminus (where residues 63 and 106 have been replaced with
CC glycine). E7B6Pentm is representative of fusion proteins of the
CC invention. The specification describes human papillomavirus E6 and E7
CC polypeptides, where the E7 polypeptide has mutations at any one or more
CC of the amino acids corresponding to amino acids 24, 26 or 91 of the
CC sequence given in ADO44073 and the E6 polypeptide has no mutations or has
CC mutations at any one or more of the amino acids corresponding to amino
CC acids 63 or 106 of the sequence given in ADO44072. The polypeptides of
CC the invention are useful for treating or preventing human papillomavirus
CC (HPV)-associated cancers, such as cervical cancer. The fusion proteins
CC and nucleic acids encoding the fusion proteins are useful for generating
CC immune responses against HPV. They are also useful for treating lower
CC gastrointestinal tract cancers, e.g. anal cancer, and other cancers of
CC the reproductive system, including penile and vulvar cancer.
XX
XX Sequence 747 BP; 246 A; 132 C; 171 G; 198 T; 0 U; 0 Other;
SQ
Query Match 100.0%; Score 28; DB 12; Length 747;
Best Local Similarity 100.0%; Pred. No. 0.036;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```


Db 556 GAACAGCAATACAAACCGTGTGTG 583

RESULT 24
ADRA7004

ID ADRA7004 standard; DNA; 768 BP.

AC ADRA7004;

DT 18-NOV-2004 (first entry)

DE Human papillomavirus type 16 E7 gene for Dengue virus vaccine.

KM de; gene; cytostatic; virucide; dengue virus; recombinant replicon;
deletion; prem protein; C protein; NSI protein signal; vaccine;
cervical cancer; viral disease; antigen; dendritic cell; immune response;
human papillomavirus.

OS Human papillomavirus type 16.

PH Key Location/Qualifiers

FT CDS 1..768

FT /tag= a

FT /product= "HPV-16 E7 protein"

PN WO2004072274-A1.

PD 26-AUG-2004.

PP 30-JAN-2004; 2004WO-CN000088.

PR 30-JAN-2003; 2003CN-00115272.

PR 30-JAN-2003; 2003CN-00115273.

PA (SHAN-) SHANGHAI TENGGEN BIOMEDICAL CO LTD.

PA (TENG-) TENGGEN BIOMEDICAL CO.

PA (BEIJ-) BEIJING ORIENTAL TENGGEN TECHNOLOGY DEV C.

PI Pang X;

PI WPI; 2004-625870/60.

PI P-PSDB; ADRA7005.

PI GENBANK; AF486352.

PT virus-like particle vaccines containing dengue virus recombinant replicon

PT as core for carrier, applicable in preventives or/and remedies for tumors

PT like cervical cancer and viral diseases.

XX Example 2; SEQ ID NO 1; 38pp; Chinese.

XX A dengue virus recombinant replicon has a deletion of the complete coding

XX sequence for prem protein of dengue virus and also includes elements of

XX e.g. the non-coding region in the whole of the 5'-end, the coding region

XX of the front 20 amino acids in the C protein, and the coding region of

XX NSI protein signal; coding regions of all non-structural proteins. The

XX obtained vaccines are useful in producing preventives or/and remedies for

XX efficiently express antigen in infected cells, which is because dengue

XX virus can infect dendritic cells, and can effectively present antigen to

XX provide immunity effect. Different types of dengue virus can be used to

XX repeatedly produce efficient immune response thereby strengthening the

XX body's immune system against the pathogen that contains such antigen.

XX Human papillomavirus (HPV) vaccines were prepared by using a gene-

XX expressing system using of the full-length dengue virus cDNA clone

XX (P8S/PLD2). The recombinant virus vectors were transfected into baby

XX hamster kidney (BHK) cells to enable the screening of BHK-21 Tet-off cell

XX lines. This sequence corresponds to the HPV type 16 E7 gene used as the

XX gene of interest in the recombinant replicon of the invention.

XX Sequence 768 BP; 255 A; 139 C; 172 G; 202 T; 0 U; 0 Other;

XX Query Match 100.0%; Score 28; DB 13; Length 768;

Best Local Similarity 100.0%; Pred. No. 0.036;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAACAGCAATACAAACCGTGTGTG 28
Db 580 GAACAGCAATACAAACCGTGTGTG 607

RESULT 25

ID AEF40156 standard; DNA; 768 BP.

AC AEF40156;

DT 23-MAR-2006 (first entry)

DE Human papillomavirus 16 (HPV-16) E7-E6 oncogene.

KM Vaccine; virus-like particle; replicon; therapeutic; cancer; cytostatic;

KM neoplasm; viral infection; virucide; infection; oncogene;

KM coding sequence; de.

OS Human papillomavirus type 16.

PH Key Location/Qualifiers

FT CDS 1..768

FT /tag= a

FT /partial

FT /product= "Human papillomavirus 16 (HPV-16) E7-E6

FT oncoprotein"

FT /note= "No stop codon"

PN US2006018928-A1.

PD 26-JAN-2006.

PP 29-JUL-2005; 2005US-00192923.

PR 30-JAN-2003; 2003CN-00115272.

PR 30-JAN-2003; 2003CN-00115273.

PR 30-JAN-2004; 2004WO-00072274.

PA (PANG/) PANG X.

PA Pang X;

PI WPI; 2006-109169/11.

PI P-PSDB; AEF40157.

PI GENBANK; AF486352, AF469197, AF472508.

PT New recombinant DEN replicons with a deletion of prem, useful for

PT producing a drug for the prophylaxis and treatment of cancer or viral

PT infection.

XX Example 2; SEQ ID NO 1; 24pp; English.

XX The present invention provides a virus-like particle (VLP) vaccine which

XX contains dengue virus (DEN) recombinant replicon as its core. The DEN

XX replicon contains exogenous nucleotide sequences such as human

XX papillomavirus (HPV) antigen proteins, immune regulators or combination

XX of HPV antigen and immune regulators. The invention is useful for

XX producing a drug for the prophylaxis and treatment of cancer or viral

XX infection. The present sequence is a human papillomavirus oncogene.

XX Sequence 768 BP; 255 A; 139 C; 172 G; 202 T; 0 U; 0 Other;

XX Query Match 100.0%; Score 28; DB 15; Length 768;

XX Best Local Similarity 100.0%; Pred. No. 0.036;

XX Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX Qy 1 GAACAGCAATACAAACCGTGTGTG 28

XX Db 580 GAACAGCAATACAAACCGTGTGTG 607

```

RESULT 26
AAT14663
ID AAT14663 standard; DNA; 776 BP.
XX
AC AAT14663;
XX
DT 27-AUG-2003 (revised)
DT 25-MAR-2003 (revised)
DT 10-OCT-1996 (first entry)
XX
DE B6/E7 region of Human Papilloma Virus 16 (HPV 16).
XX
XX Human papilloma virus; HPV; detection; cervical cancer; amplification;
XX hybridisation; diagnosis; transformed cell; B6; E7; ss.
XX
OS Human papillomavirus type 16.
XX
Key Location/Qualifiers
FH 1. .24
FT primer_bind /*tag= a
FT /note= "Primer BB113 binding site."
FT m18c_binding /tag= b
FT 30. .55 /note= "Primer H16-58 binding site."
FT 37. .57 /*tag= c
FT primer_bind /*tag= "Primer BB4 binding site."
FT 454. .474 /tag= d
FT /note= "Primer BB114 binding site."
FT 480. .503 /*tag= e
FT /note= "Primer BB11 binding site."
FT 591. .621 /*tag= f
FT primer_bind /note= "Primer BB109 binding site."
FT 658. .681 /*tag= g
FT primer_bind /note= "Primer H16-686 binding site."
FT 660. .683 /*tag= h
FT /note= "Primer BB112 binding site."
FT 715. .738 /*tag= i
FT primer_bind /note= "Primer H16-743 binding site."
FT 745. .768 /*tag= j
FT primer_bind /note= "Primer H16-773 binding site."
XX
XX US5506105-A.
XX
XX 09-APR-1996.
XX
XX 22-MAR-1994; 94US-00216233.
XX
XX 10-DEC-1991; 91US-00808456.
XX
XX (DADA-) DADA INT INC.
XX
XX Haydock PV;
XX
XX WPI; 1996-200273/20.
XX
XX Detection of low copy number intracellular markers - by 3SR amplification
XX of target RNA in fixed cells then hybridisation with labelled probe, for
XX detecting human papilloma virus in cervical cells.
XX
XX Example 2; Fig 3; 21pp; English.
XX
XX An in situ hybridisation assay for detecting an intracellular marker of
XX low copy number in cells comprises fixing the cells to a support using
CC

```

```

CC paraformaldehyde; treating the cells with a protease to permeabilise them
CC without altering morphology; adding amplification reagents; incubating
CC the cells at below fifty degrees celsius to perform amplification by self
CC -sustained sequence replication; adding a labelled probe complementary to
CC the region between the primers; washing cells to remove unhybridised
CC probe and then detecting the labelled probe. The B6/E7 region of human
CC papilloma virus (HPV) 16 is used especially to detect mRNA being
CC transcribed from this region which becomes active in transformed cells.
CC The method can be used for the early diagnosis of cervical cancer.
CC Primers used to amplify fragments of the B6/E7 region are described in
CC AAT14664-T14674. (Updated on 25-MAR-2003 to correct PF field.) (Updated
CC on 27-AUG-2003 to correct OS field.)
XX
XX SQ Sequence 776 BP; 260 A; 139 C; 169 G; 208 T; 0 U; 0 Other;
XX
XX Query Match 100.0%; Score 28; DB 2; Length 776;
XX Best Local Similarity 100.0%; Pred. No. 0.036;
XX Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
Qy 1 GAACAGCAATACACAAACCGTTGTGTG 28
Db 286 GAACAGCAATACACAAACCGTTGTGTG 313
XX
RESULT 27
AAX89756
ID AAX89756 standard; DNA; 779 BP.
XX
XX AAX89756;
XX
XX 27-AUG-2003 (revised)
XX DT 12-OCT-1999 (first entry)
XX
XX DE Probe sequence for HPV 16 B6/E7.
XX
XX KW human papilloma virus; infection; gene expression; probe; detection;
XX assay; cancer; virus; HPV; ss.
XX
XX OS Synthetic.
XX OS Human papillomavirus.
XX PN MO9929890-A2.
XX
XX 17-JUN-1999.
XX
XX 11-DEC-1998; 98MO-US026447.
XX
XX 12-DEC-1997; 97US-0069426P.
XX PR 05-JAN-1998; 98US-0070486P.
XX PR 17-APR-1998; 98US-0082167P.
XX
XX (DIGE-) DIGENE CORP.
XX
XX Lorincz AT;
XX
XX WPI; 1999-443850/37.
XX
XX New method for assessing Human Papilloma Virus (HPV) infection by
XX comparison of gene expression levels.
XX
XX Disclosure; Fig 5; 35pp; English.
XX
XX This nucleotide probe is specific for the HPV16 Human Papilloma Virus
XX (HPV) gene. The degree of HPV infection can be assessed, by measuring the
XX levels of expression of genes involved in the diseased state, and
XX comparing the expression to each other or to reference genes. (Updated on
XX 27-AUG-2003 to correct OS field.)
XX
XX SQ Sequence 779 BP; 259 A; 139 C; 170 G; 211 T; 0 U; 0 Other;
XX
XX Query Match 100.0%; Score 28; DB 2; Length 779;
XX Best Local Similarity 100.0%; Pred. No. 0.036;
XX Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```



```

Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      1 GAACGCAATATACAAACCGTTGTG 28
      |||||
Db      286 GAACGCAATATACAAACCGTTGTG 313

RESULT 30
AAK78792
ID      AAK78792 standard; DNA; 822 BP.
AC      AAK78792;
XX
XX
DT      06-SEP-1999 (first entry)
XX
XX
DE      HPV fusion protein D1/3-B6-His/HPV16 DNA.
XX
XX
KM      Fusion protein; E6 protein; E7 protein; E6/E7; immunomodulator; tumour;
KW      Immunological fusion partner; CpG oligonucleotide; immune response;
KW      HPV antigen; prevention; treatment; ss.
XX
XX
OS      Synthetic.
OS      Human papillomavirus.
XX
XX
PN      WO933868-A2.
XX
XX
PD      08-JUL-1999.
XX
XX
PF      18-DEC-1998; 98WO-EP008563.
XX
XX
PR      24-DEC-1997; 97GB-00027262.
XX
XX
PA      (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.
XX
PI      Dalemans WLJ, Gerard CMG;
XX
XX
DR      WPI; 1999-405485/34.
DR      P-PDB; AAY25376.
XX
XX
PT      Composition comprising an E6, E7 or E6/E7 fusion protein from HPV to
PT      induce immune response to HPV.
XX
XX
PS      Example II; Page 48; 62pp; English.
XX
XX
CC      AAK78791-X78801 represent nucleic acid sequences which encode novel
CC      constructs comprising an E6 or E7 protein or E6/E7 fusion protein from
CC      HPV (represented in AAY25375-Y25386). These constructs are optionally
CC      linked to an immunological fusion partner and an immunomodulatory CpG
CC      oligonucleotide. The products of the invention can be used to induce an
CC      immune response in a patient to an HPV antigen. They can also be used for
CC      preventing or treating HPV induced tumours
XX
XX
SQ      Sequence 822 BP; 277 A; 147 C; 168 G; 230 T; 0 U; 0 Other;
XX
Query Match      100.0%; Score 28; DB 2; Length 822;
Best Local Similarity 100.0%; Pred. No. 0.037;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      1 GAACGCAATATACAAACCGTTGTG 28
      |||||
Db      604 GAACGCAATATACAAACCGTTGTG 631

RESULT 31
AAK29781
ID      AAK29781 standard; DNA; 822 BP.
AC      AAK29781;
XX
XX
DT      17-OCT-2003 (revised)
DT      22-JUN-1999 (first entry)
XX
XX
DE      Prot. D1/3-B6-His/HPV16 coding sequence.

```

```

XX
XX
KM      Chimeric; E6; E7; fusion protein; protein D; vaccine; immunotherapy;
KW      tumour; lesion; benign; malignant; virus; infection; ss.
XX
XX
OS      Human papillomavirus.
OS      Haemophilus influenzae.
OS      Chimeric.
XX
XX
PN      WO9910375-A2.
XX
XX
PD      04-MAR-1999.
XX
XX
PF      17-AUG-1998; 98WO-EP005285.
XX
XX
PR      22-AUG-1997; 97GB-00017953.
XX
XX
PA      (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.
XX
PI      Bruck C, Cabazon Silva T, Delisse ABF, Gerard CMG;
PI      Lombardo-Bencheikh A;
XX
XX
DR      WPI; 1999-190587/16.
XX
XX
PT      Human Papilloma Virus (HPV) fusion proteins - useful in vaccines for
PT      treatment or prophylaxis of HPV induced lesions.
XX
XX
PS      Disclosure; Fig 3; 95pp; English.
XX
XX
CC      This sequence represents the coding region for a chimeric E6 or E7
CC      protein or E6/E7 fusion protein from Human papillomavirus (HPV) linked to
CC      an immunological fusion partner. In this case, a fragment of the
CC      Haemophilus influenzae B protein D. The sequence also contains a
CC      histidine tag at the C-terminus of the encoded protein. The protein can
CC      be used in a vaccine, for immuno-therapeutically treating HPV induced
CC      tumour lesions (benign or malignant) and preventing HPV viral infection.
CC      (updated on 17-OCT-2003 to standardise OS field)
XX
XX
SQ      Sequence 822 BP; 277 A; 147 C; 168 G; 230 T; 0 U; 0 Other;
XX
Query Match      100.0%; Score 28; DB 2; Length 822;
Best Local Similarity 100.0%; Pred. No. 0.037;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      1 GAACGCAATATACAAACCGTTGTG 28
      |||||
Db      604 GAACGCAATATACAAACCGTTGTG 631

RESULT 32
AED52633
ID      AED52633 standard; DNA; 822 BP.
AC      AED52633;
XX
XX
DT      29-DEC-2005 (first entry)
XX
XX
DE      Fusion protein D1/3-B6-His/HPV16, DNA.
XX
XX
KM      Fusion protein; vaccine; papilloma; cytostatic; papillomavirus infection;
KW      virucide; uterine cervix tumor; E6; del gene; D protein.
XX
XX
OS      Haemophilus influenzae; strain 772.
OS      Human papillomavirus type 16.
OS      Synthetic.
OS      Chimeric.
XX
XX
FH      Key
FT      CDS
FT      CDS
FT      CDS
XX
XX
DE      IN9801903-14.

```

PD 04-MAR-2005.
 XX
 XX 24-AUG-1998; 98IN-CH001903.
 PF
 XX 22-AUG-1997; 97EP-00179535.
 PR
 XX (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
 PA
 XX Tyrell AMR;
 PI
 XX WPI; 2005-557648/57.
 DR
 XX P-PSDB; AED52634.
 DR
 XX
 XX Vaccine.
 PT
 PS Example 4; Fig 3; 96pp; English.
 XX
 CC The invention relates to human Papilloma virus (HPV) fusion proteins,
 CC linked to an immunological fusion partner that provides T helper epitopes
 CC to the HPV antigen. Vaccine formulations comprising the fusion proteins
 CC are useful in the treatment or prophylaxis of HPV induced lesions
 CC (papillomas, warts and cervical cancer). HPV 16 (and HPV 18) E6 and E7
 CC proteins (singly, as an E6-E7 fusion or mutated) were fused to either
 CC Haemophilus influenzae D protein (20-127), the C-terminus of
 CC Streptococcus pneumoniae Lyta protein (CLYTA) or thiorodoxin. The present
 CC sequence encodes an HPV-H. influenzae D protein, fusion protein of the
 CC invention.
 CC
 XX Sequence 822 BP; 277 A; 147 C; 168 G; 230 T; 0 U; 0 Other;
 SQ
 XX
 XX Query Match 100.0%; Score 28; DB 14; Length 822;
 Best Local Similarity 100.0%; Pred. No. 0.037;
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 GAACAGCAATACAAACACCGTTGTGTG 28
 Db 604 GAACAGCAATACAAACACCGTTGTGTG 631
 XX
 XX
 XX RESULT 33
 AAX78795
 ID AAX78795 standard; DNA; 879 BP.
 XX
 AC AAX78795;
 XX
 DT 06-SEP-1999 (first entry)
 XX
 DE HPV fusion protein CLYTA-E6-His/HPV16 DNA.
 XX
 KW Fusion protein; E6 protein; E7 protein; E6/E7; immunomodulator; tumour;
 KW immunological fusion partner; Cpg oligonucleotide; immune response;
 KW HIV antigen; prevention; treatment; ss.
 XX
 XX Synthetic.
 OS Human papillomavirus.
 OS
 PN WO933868-A2.
 PN
 PD 08-JUL-1999.
 PD
 PF 18-DEC-1998; 98MO-EP008563.
 PF
 PR 24-DEC-1997; 97GB-00027262.
 PR
 XX (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
 PA
 PI Dalemans WLJ, Gerard CMG;
 PI
 CC WPI; 1999-405485/34.
 CC
 DR P-PSDB; AAY25379.
 DR
 XX Composition comprising an E6, E7 or E6/E7 fusion protein from HPV to
 PT induce immune response to HPV.
 PT

XX
 PS Example VI; Page 52; 62pp; English.
 XX
 CC AAX78791-X78801 represent nucleic acid sequences which encode novel
 CC constructs comprising an E6 or E7 protein or E6/E7 fusion protein from
 CC HPV (represented in AAY25375-Y25386). These constructs are optionally
 CC linked to an immunological fusion partner and an immunomodulatory Cpg
 CC oligonucleotide. The products of the invention can be used to induce an
 CC immune response in a patient to an HPV antigen. They can also be used for
 CC preventing or treating HPV induced tumours
 CC
 XX Sequence 879 BP; 289 A; 168 C; 209 G; 213 T; 0 U; 0 Other;
 SQ
 XX
 XX Query Match 100.0%; Score 28; DB 2; Length 879;
 Best Local Similarity 100.0%; Pred. No. 0.037;
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 GAACAGCAATACAAACACCGTTGTGTG 28
 Db 661 GAACAGCAATACAAACACCGTTGTGTG 688
 XX
 XX
 XX RESULT 34
 AAX29784
 ID AAX29784 standard; DNA; 879 BP.
 XX
 AC AAX29784;
 XX
 DT 17-OCT-2003 (revised)
 DT 22-JUN-1999 (first entry)
 XX
 DE CLYTA-E6-His coding sequence.
 XX
 KW Chimeric; E6; E7; fusion protein; CLYTA; vaccine; immunotherapy; tumour;
 KW lesion; benign; malignant; virus; infection; ss.
 XX
 XX Human papillomavirus.
 OS Streptococcus pneumoniae.
 OS Chimeric.
 OS
 PN WO9910375-A2.
 PN
 PD 04-MAR-1999.
 PD
 PF 17-AUG-1998; 98MO-EP005285.
 PF
 PR 22-AUG-1997; 97GB-00017953.
 PR
 XX (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
 PA
 PI Bruck C, Cabezon Silva T, Delisse ABF, Gerard CMG;
 PI Lombardo-Bencheikh A;
 PI
 CC WPI; 1999-190587/16.
 CC
 DR P-PSDB; AAY02635.
 DR
 XX Human Papilloma Virus (HPV) fusion proteins - useful in vaccines for
 PT treatment or prophylaxis of HPV induced lesions.
 PT
 PS Disclosure; Fig 10; 95pp; English.
 PS
 XX This sequence represents the coding region for a chimeric E6 or E7
 CC protein or E6/E7 fusion protein from human papillomavirus (HPV) linked to
 CC an immunological fusion partner, in this case, a fragment of the
 CC Streptococcus pneumoniae CLYTA protein. The sequence also contains a
 CC histidine tag at the C-terminus of the encoded protein. The protein can
 CC be used in a vaccine, for immuno-therapeutically treating HPV induced
 CC tumour lesions (benign or malignant) and preventing HPV viral infection.
 CC (Updated on 17-OCT-2003 to standardise OS field)
 CC
 SQ Sequence 879 BP; 289 A; 168 C; 209 G; 213 T; 0 U; 0 Other;
 XX
 XX Query Match 100.0%; Score 28; DB 2; Length 879;
 PT


```

ID AAD35101 standard; DNA; 939 BP.
XX
AC AAD35101;
XX
DT 25-JUL-2002 (first entry)
XX
DE Human papillomavirus enhE6,7 construct DNA.
XX
KM Alphavirus vector system; human papilloma virus; HPV; cervical cancer;
XX therapy; vaccine; virucide; cytostatic; ds.
XX
OS Human papillomavirus.
XX
PN BP195438-A1.
XX
PD 10-APR-2002.
XX
PF 06-OCT-2000; 2000BP-00203472.
XX
PR 06-OCT-2000; 2000BP-00203472.
XX
PA (UYGR-) RIJXSUNITV GRONINGEN.
XX
PI Regts DG, Wilschut JC, Holtrop M, Daemen CAH;
XX
DR WPI; 2002-354156/39.
XX
PT New alphavirus system, useful for genetic immunization against cervical
XX cancer, comprises papilloma virus nucleic acid.
XX
PS Example 2; Fig 19; 45pp; English.
XX
CC The present invention relates to an alphavirus vector system comprising
XX nucleic acid derived from a human papilloma virus (HPV). The invention or
XX cells containing it, are used in treatment and prevention of cervical
XX cancer, particularly as a vaccine. By selecting the nucleic acid that
XX encode E6/E7 proteins without ability to bind to pRb and p53, the risk
XX that cells infected with the alphavirus vector system may become
XX oncogenic is avoided (contrast use of other viral vectors). The present
XX sequence is Human papillomavirus enhE6,7 construct DNA
XX
SQ Sequence 939 BP; 284 A; 197 C; 217 G; 241 T; 0 U; 0 Other;

Query Match 100.0%; Score 28; DB 6; Length 939;
Best Local Similarity 100.0%; Pred. No. 0.037;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAACAGCATATCAACAACCGTTGTGTG 28
DB 450 GAACAGCATATCAACAACCGTTGTGTG 477

RESULT 38
AAQ08627
ID AAQ08627 standard; DNA; 1000 BP.
XX
AC AAQ08627;
XX
DT 24-OCT-2003 (revised)
XX 21-APR-1994 (first entry)
XX
DE HPV-16 fragment.
XX
KM HPV-18; HPV-16; amplification; primer; polymerase chain reaction; PCR;
XX ss.
XX
OS Human papillomavirus type 16.
XX
PN Key Location/Qualifiers
XX misc_binding 198..207
XX /tag= a
XX /note= "primer (AAQ08628) binding site"
XX
FT misc_binding 601..620

```

```

FT /*tag= b
FT /note= "primer (AAQ08629) binding site"
FT misc_binding 658..677
FT /*tag= c
FT /note= "primer (AAQ08630) binding site"
XX
PN DE3838269-A.
XX
PD 17-MAY-1990.
XX
PF 11-NOV-1988; 88DE-03838269.
XX
PR 11-NOV-1988; 88DE-03838269.
XX
PA (BEHW ) BEHRINGERKE AG.
XX
PI Cerutti P, Whitcomb J, Zijlstra J, Devillers EM;
XX
DR WPI; 1990-156905/21.
XX
PT Detection of human papilloma virus - by DNA amplification and analysis.
XX
PS Example 4b; Page 4; 11pp; German.
XX
CC Example 4b describes the results of the amplification of HPV-16 DNA by
XX PCR using primers. (Updated on 24-OCT-2003 to standardise OS field)
XX
SQ Sequence 1000 BP; 340 A; 171 C; 230 G; 259 T; 0 U; 0 Other;

Query Match 100.0%; Score 28; DB 2; Length 1000;
Best Local Similarity 100.0%; Pred. No. 0.037;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAACAGCATATCAACAACCGTTGTGTG 28
DB 368 GAACAGCATATCAACAACCGTTGTGTG 395

RESULT 39
AAN91784
ID AAN91784 standard; DNA; 1005 BP.
XX
AC AAN91784;
XX
DT 27-AUG-2003 (revised)
XX 25-MAR-2003 (revised)
XX 16-MAR-1990 (first entry)
XX
DE DNA probe complementary to human papilloma virus type 16.
XX
KM Cervical cancer.
XX
OS Human papillomavirus type 16.
XX
PN WO8909940-A.
XX
PD 19-OCT-1989.
XX
PF 04-APR-1989; 89WO-US001318.
XX
PR 04-APR-1988; 88US-00177404.
XX 31-MAR-1989; 89US-00330381.
XX
PA (ONCO-) ONCOR INC.
XX
PI George AL, Groff DE;
XX
DR WPI; 1989-324314/44.
XX
PT Rapid detection of specific human papilloma virus genotypes - by
XX hybridisation of DNA digest with new labelled nucleic acid probes.
XX
PS Claim 40; Page 46; 81pp; English.

```

XX Obtd. by cutting HPV16 with BamHI and PvuII. The patent describes probes
CC (DNA or RNA) and their complements capable of detecting one or a
CC combination of HPV types 6, 11, 16, 18, 31, 33 and 35. (Updated on 25-MAR
CC -2003 to correct PR field.) (Updated on 27-AUG-2003 to correct OS field.)
XX

Sequence 1005 BP; 325 A; 182 C; 190 G; 308 T; 0 U; 0 Other;

Query Match 100.0%; Score 28; DB 1; Length 1005;
Best Local Similarity 100.0%; Pred. No. 0.037;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAACAGCAATACAAACCGTGTGTG 28
Db 820 GAACAGCAATACAAACCGTGTGTG 847

RESULT 40

AAx78793
ID AAX78793 standard; DNA; 1116 BP.

XX AAX78793;

XX 06-SEP-1999 (first entry)

XX HPV fusion protein D1/3-B6-E7-His/HPV16 DNA.

XX Fusion protein; E6 protein; E7 protein; E6/E7; immunomodulator; tumour;
XX immunological fusion partner; CpG oligonucleotide; immune response;
XX HPV antigen; prevention; treatment; ss.

XX Synthetic.
XX Human papillomavirus.

XX WO933868-A2.

XX 08-JUL-1999.

XX 18-DEC-1998; 98WO-EP008563.

XX 24-DEC-1997; 97GB-00027262.

XX (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.

XX Dalemans WLF, Gerard CMG;

XX WPI; 1999-405485/34.

XX P-PsDB; AAY25377.

XX Composition comprising an E6, E7 or E6/E7 fusion protein from HPV to
XX induce immune response to HPV.

XX Example III; Page 49; 62pp; English.

XX AAX78791-X78801 represent nucleic acid sequences which encode novel
CC constructs comprising an E6 or E7 protein or E6/E7 fusion protein from
CC HPV (represented in AAY25375-Y25386). These constructs are optionally
CC linked to an immunological fusion partner and an immunomodulatory CpG
CC oligonucleotide. The products of the invention can be used to induce an
CC immune response in a patient to an HPV antigen. They can also be used for
CC preventing or treating HPV induced tumours

XX Sequence 1116 BP; 368 A; 208 C; 235 G; 305 T; 0 U; 0 Other;

Query Match 100.0%; Score 28; DB 2; Length 1116;
Best Local Similarity 100.0%; Pred. No. 0.038;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAACAGCAATACAAACCGTGTGTG 28
Db 604 GAACAGCAATACAAACCGTGTGTG 631

RESULT 41

AAx29782
ID AAX29782 standard; DNA; 1116 BP.

XX AAX29782;

XX 17-OCT-2003 (revised)

XX 22-JUN-1999 (first entry)

XX Prot. D1/3-B6-E7-His/HPV16 coding sequence.

XX Chimeric; E6; E7; fusion protein; protein D; vaccine; immunotherapy;
XX tumour; lesion; benign; malignant; virus; infection; ss.

XX Human papillomavirus.

XX Haemophilus influenzae.

XX Chimeric.

XX WO9910375-A2.

XX 04-MAR-1999.

XX 17-AUG-1998; 98WO-EP005285.

XX 22-AUG-1997; 97GB-00017953.

XX (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.

XX Bruck C, Cabezon Silva T, Delisse ABF, Gerard CMG;
XX Lombardo-Bencheikh A;

XX WPI; 1999-190587/16.

XX P-PsDB; AAY02633.

XX Human Papilloma Virus (HPV) fusion proteins - useful in vaccines for
XX treatment or prophylaxis of HPV induced lesions.

XX Disclosure; Fig 6; 95pp; English.

XX This sequence represents the coding region for a chimeric E6 or E7
CC protein or E6/E7 fusion protein from Human papillomavirus (HPV) linked to
CC an immunological fusion partner. In this case, a fragment of the
CC Haemophilus influenzae B protein D. The sequence also contains a
CC histidine tag at the C-terminus of the encoded protein. The protein can
CC be used in a vaccine, for immuno-therapeutically treating HPV induced
CC tumour lesions (benign or malignant) and preventing HPV viral infection.
CC (Updated on 17-OCT-2003 to standardise OS field)

XX Sequence 1116 BP; 368 A; 208 C; 235 G; 305 T; 0 U; 0 Other;

Query Match 100.0%; Score 28; DB 2; Length 1116;
Best Local Similarity 100.0%; Pred. No. 0.038;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAACAGCAATACAAACCGTGTGTG 28
Db 604 GAACAGCAATACAAACCGTGTGTG 631

RESULT 42

AAx52637
ID AED52637 standard; DNA; 1116 BP.

XX AED52637;

XX 29-DEC-2005 (first entry)

XX Fusion protein D1/3-B6-E7-His/HPV16, DNA.

XX Fusion protein; vaccine; papilloma; cytostatic; papillomavirus infection;
XX virucide; uterine cervix tumor; E7; E6; de; gene; D protein.

XX Haemophilus influenzae; strain 772.

OS	Human papillomavirus type 16.
OS	Synthetic.
OS	Chimeric.
XX	
FX	Key
FT	Location/Qualifiers
FT	1..1116
FT	/tag= a
FT	/product= "Fusion protein DI/3-E6-E7-His/HPV16"
XX	
PN	IN9801903-14.
XX	
PD	04-MAR-2005.
XX	
PE	24-AUG-1998;
XX	98IN-CH001903.
PR	22-AUG-1997;
XX	97EP-00179535.
PA	(SMIK) SMTKLINE BEECHAM BIOLOGICALS.
XX	
P1	Tyrrell AWR;
DR	WPI; 2005-557648/57.
DR	P-P5DB; AED52638.
XX	
PT	Vaccine.
XX	
PS	Example 6; Fig 6; 96dp; English.
XX	
CC	The invention relates to human Papilloma virus (HPV) fusion proteins,
CC	linked to an immunological fusion partner that provides T helper epitopes
CC	to the HPV antigen. Vaccine formulations comprising the fusion proteins
CC	are useful in the treatment or prophylaxis of HPV induced lesions
CC	(papillomas, warts and cervical cancer). HPV 16 and HPV 18) B6 and E7
CC	proteins (singly, as an B6-E7 fusion or mutated) were fused to either
CC	Haemophilus influenzae D protein (20-127), the C-terminus of
CC	Streptococcus pneumoniae LytA protein (Clyta) or chlorodoxin. The present
CC	sequence encodes an HPV-H. influenzae D protein, fusion protein of the
XX	invention.
XX	
SQ	Sequence 1116 BP; 368 A; 208 C; 235 G; 305 T; 0 U; 0 Other;
QY	Query Match 100.0%; Score 28; DB 14; Length 1116;
Db	Best Local Similarity 100.0%; Pred. No. 0.038;
	Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
	1 GAACGCAATTCACACAAACCCTTGTGTG 28
	604 GAACGCAATTCACACAAACCCTTGTGTG 631
RESULT 43	
AAX78797	
ID	AAX78797 standard; DNA; 1173 BP.
XX	
AC	AAX78797;
XX	
DT	06-SEP-1999 (first entry)
XX	
DE	HPV fusion protein CLYTA-E6E7-His/HPV16 DNA.
XX	
KW	Fusion protein; B6 protein; E7 protein; B6/E7; immunomodulator; tumour;
KW	immunological fusion partner; CpG oligonucleotide; immune response;
KW	HPV antigen; prevention; treatment; ss.
XX	
OS	Synthetic.
OS	Human papillomavirus.
XX	
PN	MO9933868-A2.
XX	
PD	08-JUL-1999.
XX	
PF	18-DEC-1998; 98WO-EP008563.
XX	

```

PR 24-DEC-1997; 97GB-00027262.
XX
XX (SMIK ) SMITHLINE BEECHAM BIOLOGICALS.
XX
XX Dalemans WJ, Gerard CMG;
XX
XX WPI, 1999-405485/34.
XX P-PSDB; AAY25381.
XX
XX Composition comprising an E6, E7 or E6/E7 fusion protein from HPV to
XX induce immune response to HPV.
XX
XX Example VIII, Page 54-55; 62pp; English.
XX
XX AAX78791-778801 represent nucleic acid sequences which encode novel
XX constructs comprising an E6 or E7 protein or E6/E7 fusion protein from
XX HPV (represented in AAY25375-Y25386). These constructs are optionally
XX linked to an immunological fusion partner and an immunomodulatory CpG
XX oligonucleotide. The products of the invention can be used to induce an
XX immune response in a patient to an HPV antigen. They can also be used for
XX preventing or treating HPV induced tumours
XX
XX Sequence 1173 BP, 380 A; 229 C; 276 G; 288 T; 0 U; 0 Other;
XX
XX Query Match 100.0%; Score 28; DB 2; Length 1173;
XX Best Local Similarity 100.0%; Pred. No. 0.038;
XX Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0,
XX
XX 1 GAACGCAATTCACAAACCGTTGTG 28
XX |||||||
XX 661 GAACGCAATTCACAAACCGTTGTG 688
XX
XX RESULT 44
XX AAX23786
XX ID AAX29786 standard; DNA; 1173 BP.
XX
XX AAX29786;
XX
XX 17-OCT-2003 (revised)
XX DT 22-JUN-1999 (first entry)
XX
XX CLYTA-E6E7-His coding sequence.
XX
XX Chimeric; E6; E7, fusion protein; CLYTA; vaccine; immunotherapy; tumour;
XX lesion; benign; malignant; virus; infection; ss.
XX Human papillomavirus.
XX OS Streptococcus pneumoniae.
XX OS Chimeric.
XX
XX WO9910375-A2.
XX
XX 04-MAR-1999.
XX
XX 17-AUG-1998; 98WO-EP005285.
XX
XX 22-AUG-1997; 97GB-00017953.
XX
XX (SMIK ) SMITHLINE BEECHAM BIOLOGICALS.
XX
XX Bruck C, Cabezon Silva T, Delisse AEF, Gerard CMG;
XX Lombardo-Benckel A;
XX
XX WPI, 1999-190587/16.
XX DR P-PSDB; AAY02637.
XX
XX Human Papilloma Virus (HPV) fusion proteins - useful in vaccines for
XX treatment or prophylaxis of HPV induced lesions.
XX
XX Disclosure, Fig 14; 95pp; English.
XX
XX This sequence represents the coding region for a chimeric E6 or E7

```

CC protein or E6/E7 fusion protein from Human papillomavirus (HPV) linked to
CC an immunological fusion partner, in this case, a fragment of the
CC Streptococcus pneumoniae CLYTA protein. The sequence also contains a
CC histidine tag at the C-terminus of the encoded protein. The protein can
CC be used in a vaccine, for immuno-therapeutically treating HPV induced
CC tumour lesions (benign or malignant) and preventing HPV viral infection.
CC (Updated on 17-Oct-2003 to standardise OS field)

XX
SQ Sequence 1173 BP; 380 A; 229 C; 276 G; 288 T; 0 U; 0 Other;

Query Match 100.0%; Score 28; DB 2; Length 1173;
Best Local Similarity 100.0%; Pred. No. 0.038;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAACAGCAATACAAACCGTTGTGTG 28
|||||
Db 661 GAACAGCAATACAAACCGTTGTGTG 688

RESULT 45
AED52645
ID AED52645 standard; DNA; 1173 BP.

XX AED52645;

XX 29-DEC-2005 (first entry)

XX Fusion protein cLYTA-B6-E7-His/HPV16, DNA.

XX Fusion protein; vaccine; papilloma; cytosolic; papillomavirus infection;
XX virulence; uterine cervix tumor; E7; E6; de; gene; LyTA.

XX Streptococcus pneumoniae.

XX Human papillomavirus type 16.

XX Synthetic.

XX Chimeric.

XX Key Location/Qualifiers
FT 1..1173
FT /tag= a
FT /product= "Fusion protein cLYTA-B6-E7-His/HPV16"

XX IN9801903-14.

XX 04-MAR-2005.

XX 24-AUG-1998; 98IN-CH001903.

XX 22-AUG-1997; 97EP-00179535.

XX (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.

XX Tytrell AMR;

XX WPI; 2005-557648/57.

XX P-PSDB; AED52646.

XX Vaccine.

XX Example 12; Fig 14; 96pp; English.

CC The invention relates to human Papilloma virus (HPV) fusion proteins,
CC linked to an immunological fusion partner that provides T helper epitopes
CC to the HPV antigen. Vaccine formulations comprising the fusion proteins
CC are useful in the treatment or prophylaxis of HPV induced lesions
CC (papillomas, warts and cervical cancer). HPV 16 (and HPV 18) E6 and E7
CC proteins (singly, as an E6-E7 fusion or mutated) were fused to either
CC Haemophilus influenzae D protein (20-127), the C-terminus of
CC Streptococcus pneumoniae LyTA protein (cLYTA) or thiorodoxin. The present
CC sequence encodes an HPV-LyTA, fusion protein of the invention.

XX Sequence 1173 BP; 380 A; 229 C; 276 G; 288 T; 0 U; 0 Other;

Query Match 100.0%; Score 28; DB 14; Length 1173;
Best Local Similarity 100.0%; Pred. No. 0.038;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAACAGCAATACAAACCGTTGTGTG 28
|||||
Db 661 GAACAGCAATACAAACCGTTGTGTG 688

RESULT 46

AAFS5127
ID AAF55127 standard; DNA; 7840 BP.

XX AAF55127;

XX 29-MAY-2001 (first entry)

XX Nucleotide sequence of the vector pRetroFF-B6E7.

XX Stem cell; gene therapy; cell therapy; stem cell disorder; ss.

XX Synthetic.

XX Key Location/Qualifiers

XX misc_feature

XX /tag= d
XX /note= "this base is given as N because it was illegible
XX in the figure"

XX /tag= e
XX /note= "this base is given as N because it was illegible
XX in the figure"

XX /tag= f
XX /note= "this base is given as N because it was illegible
XX in the figure"

XX /tag= g
XX /note= "this base is given as N because it was illegible
XX in the figure"

XX /tag= h
XX /note= "this base is given as N because it was illegible
XX in the figure"

XX /tag= i
XX /note= "this base is given as N because it was illegible
XX in the figure"

XX /tag= j
XX /note= "this base is given as N because it was illegible
XX in the figure"

XX /tag= k
XX /note= "this base is given as N because it was illegible
XX in the figure"

XX /tag= l
XX /note= "this base is given as N because it was illegible
XX in the figure"

XX /tag= m
XX /note= "this base is given as N because it was illegible
XX in the figure"

XX /tag= n
XX /note= "this base is given as N because it was illegible
XX in the figure"

XX /tag= o
XX /note= "this base is given as N because it was illegible
XX in the figure"

XX /tag= p
XX /note= "this base is given as N because it was illegible
XX in the figure"

XX /tag= q
XX /note= "this base is given as N because it was illegible
XX in the figure"

XX /tag= r
XX /note= "this base is given as N because it was illegible
XX in the figure"

XX /tag= s
XX /note= "this base is given as N because it was illegible
XX in the figure"

XX /tag= t
XX /note= "this base is given as N because it was illegible
XX in the figure"

XX /tag= u
XX /note= "this base is given as N because it was illegible
XX in the figure"

XX /tag= v
XX /note= "this base is given as N because it was illegible
XX in the figure"

XX /tag= w
XX /note= "this base is given as N because it was illegible
XX in the figure"

XX /tag= x
XX /note= "this base is given as N because it was illegible
XX in the figure"

XX /tag= y
XX /note= "this base is given as N because it was illegible
XX in the figure"

XX /tag= z
XX /note= "this base is given as N because it was illegible
XX in the figure"

XX /tag= 1
XX /note= "this base is given as N because it was illegible
XX in the figure"

XX /tag= 2
XX /note= "this base is given as N because it was illegible
XX in the figure"

XX /tag= 3
XX /note= "this base is given as N because it was illegible
XX in the figure"

```
FT /tag= m
FT /note= "this base is given as N because it was illegible
FT in the figure"
FT misc_feature
FT 6100
FT /tag= n
FT /note= "this base is given as N because it was illegible
FT in the figure"
FT misc_feature
FT 6150
FT /tag= o
FT /note= "this base is given as N because it was illegible
FT in the figure"
FT misc_feature
FT 6200
FT /tag= p
FT /note= "this base is given as N because it was illegible
FT in the figure"
FT misc_feature
FT 6250
FT /tag= q
FT /note= "this base is given as N because it was illegible
FT in the figure"
FT misc_feature
FT 6300
FT /tag= r
FT /note= "this base is given as N because it was illegible
FT in the figure"
FT misc_feature
FT 6350
FT /tag= s
FT /note= "this base is given as N because it was illegible
FT in the figure"
FT misc_feature
FT 6400
FT /tag= t
FT /note= "this base is given as N because it was illegible
FT in the figure"
FT misc_feature
FT 6450
FT /tag= u
FT /note= "this base is given as N because it was illegible
FT in the figure"
FT misc_feature
FT 6500
FT /tag= v
FT /note= "this base is given as N because it was illegible
FT in the figure"
FT misc_feature
FT 6600
FT /tag= w
FT /note= "this base is given as N because it was illegible
FT in the figure"
FT misc_feature
FT 6650
FT /tag= x
FT /note= "this base is given as N because it was illegible
FT in the figure"
FT misc_feature
FT 6700
FT /tag= y
FT /note= "this base is given as N because it was illegible
FT in the figure"
FT misc_feature
FT 6750
FT /tag= z
FT /note= "this base is given as N because it was illegible
FT in the figure"
FT misc_feature
FT 6800
FT /tag= aa
FT /note= "this base is given as N because it was illegible
FT in the figure"
FT misc_feature
FT 6850
FT /tag= ab
FT /note= "this base is given as N because it was illegible
FT in the figure"
FT misc_feature
FT 6900
FT /tag= ac
FT /note= "this base is given as N because it was illegible
FT in the figure"
FT misc_feature
FT 6950
FT /tag= ad
FT /note= "this base is given as N because it was illegible
FT in the figure"
FT misc_feature
FT 7000
FT /tag= ae
FT /note= "this base is given as N because it was illegible
FT in the figure"
FT misc_feature
FT 7050
FT /tag= af
FT /note= "this base is given as N because it was illegible
FT in the figure"
FT misc_feature
FT 7100
FT /tag= ag
```

```
FT /note= "this base is given as N because it was illegible
FT in the figure"
FT misc_feature
FT 7050
FT /tag= ah
FT /note= "this base is given as N because it was illegible
FT in the figure"
FT misc_feature
FT 7100
FT /tag= ai
FT /note= "this base is given as N because it was illegible
FT in the figure"
FT misc_feature
FT 7150
FT /tag= aj
FT /note= "this base is given as N because it was illegible
FT in the figure"
FT misc_feature
FT 7200
FT /tag= ak
FT /note= "this base is given as N because it was illegible
FT in the figure"
FT misc_feature
FT 7250
FT /tag= al
FT /note= "this base is given as N because it was illegible
FT in the figure"
FT misc_feature
FT 7300
FT /tag= am
FT /note= "this base is given as N because it was illegible
FT in the figure"
FT misc_feature
FT 7350
FT /tag= an
FT /note= "this base is given as N because it was illegible
FT in the figure"
FT misc_feature
FT 7400
FT /tag= ao
FT /note= "this base is given as N because it was illegible
FT in the figure"
FT misc_feature
FT 7450
FT /tag= ap
FT /note= "this base is given as N because it was illegible
FT in the figure"
FT misc_feature
FT 7500
FT /tag= aq
FT /note= "this base is given as N because it was illegible
FT in the figure"
FT misc_feature
FT 7550
FT /tag= ar
FT /note= "this base is given as N because it was illegible
FT in the figure"
FT misc_feature
FT 7600
FT /tag= as
FT /note= "this base is given as N because it was illegible
FT in the figure"
FT misc_feature
FT 7650
FT /tag= at
FT /note= "this base is given as N because it was illegible
FT in the figure"
FT misc_feature
FT 7700
FT /tag= au
FT /note= "this base is given as N because it was illegible
FT in the figure"
FT misc_feature
FT 7750
FT /tag= av
FT /note= "this base is given as N because it was illegible
FT in the figure"
FT misc_feature
FT 7800
FT /tag= ax
FT /note= "this base is given as N because it was illegible
```

```
Query Match 100.0%; Score 28; DB 5; Length 7840;
Best Local Similarity 100.0%; Pred. No. 0.047;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GAACAGCAATPACAAACCGTGTGTG 28
Db 767 GAACAGCAATPACAAACCGTGTGTG 794
```

```

RESULT 47
AAT09847
ID AAT09847 standard; DNA; 7902 BP.
XX
AC AAT09847;
XX
DT 27-AUG-2003 (revised)
DT 02-JAN-1997 (first entry)
XX
DE Human papilloma virus 16 genomic DNA.
XX
KM Human papilloma virus; HPV; ribozyme; cleavage; E6; E7; keratinocyte;
infection; genital warts; neoplasia; cervical cancer; warts; ds.
XX
OS Human papillomavirus type 16.
PN MO95J1552-A2.
XX
PD 23-NOV-1995.
XX
PF 15-MAY-1995; 95WO-US006016.
XX
PR 13-MAY-1994; 94US-00242665.
XX
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
PA (UYNF-) UNIV NORTHERN ILLINOIS.
XX
PI Hampel A, Dipaolo J, Siwkowski A;
XX
DR WPI; 1996-010937/01.
XX
PT Synthetic ribozyme(s) for human papilloma virus (HPV) - useful in
treatment of cervical cancer and in detection of HPV.
XX
PS Disclosure; Page 25-29; 60pp; English.
XX
CC Synthetic ribozymes directed against human papilloma virus (HPV) 16 may
be used to cleave HPV or to detect HPV in human tissue. The ribozymes
cleave HPV to lower the production of the E6 and E7 proteins which play a
key role in keratinocyte transformation. They can be used for the
treatment of HPV infections which are associated with genital warts and
genital neoplasms. In particular they can be used for the treatment of
cervical cancer optionally with other agents such as cisplatin. Ribozymes
are described in AAT0306-T40315. (Updated on 27-AUG-2003 to correct OS
field.)
XX
SQ Sequence 7902 BP; 2601 A; 1376 C; 1507 G; 2418 T; 0 U; 0 Other;
XX
Query Match 100.0%; Score 28; DB 2; Length 7902;
Best local Similarity 100.0%; Pred. No. 0.047;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GAACAGCAATACAAACACCGTTGTG 28
Db 368 GAACAGCAATACAAACACCGTTGTG 395

```

```

OS Human papillomavirus type 16.
XX
XX Key Location/Qualifiers
FH TATA_signal 17..23
FT /tag= a
FT TATA_signal 65..71
FT /tag= b
FT CDS 83..559
FT /tag= c
FT /product= "E6 transforming protein"
FT /note= "E6 ORF from 65 to 559; putative"
FT mutation 350
FT /tag= d
FT /note= "Mutated from the normal T in the reference
sequence (AAT94723)"
FT CDS 562..858
FT /tag= e
FT /product= "E7 transforming protein"
FT /note= "E7 ORF 544 to 858; putative"
FT CDS 865..2813
FT /tag= f
FT /product= "E1 replication protein"
FT /note= "E1 interrupted ORF from 859 to 2813; putative"
FT CDS 2755..3852
FT /tag= g
FT /product= "E2 regulatory protein"
FT /note= "E2 ORF from 2725 to 3852; putative"
FT CDS 3332..3619
FT /tag= h
FT /product= "E4"
FT /note= "E4 ORF from 3332 to 3619; putative"
FT CDS 3863..4099
FT /tag= i
FT /product= "E5"
FT /note= "E5 ORF from 3863 to 4099; putative"
FT polyA_signal 4213..4218
FT /tag= j
FT /note= "Putative"
FT CDS 4235..5656
FT /tag= k
FT /product= "E2 minor capsid protein"
FT /note= "E2 ORF from 4133 to 5656; putative"
FT TATA_signal 4289..4295
FT /tag= l
FT CDS 5559..7154
FT /tag= m
FT /product= "E1 major capsid protein"
FT /note= "E1 ORF from 5526 to 7154; putative"
FT polyA_signal 7260..7265
FT /tag= n
XX
XX US5679509-A.
XX PN 21-OCT-1997.
XX PD 30-SEP-1994; 94US-00316239.
XX PF 28-SEP-1993; 93US-00127906.
XX PR (UYNF-) UNIV NEW MEXICO STATE.
XX PA Wheeler CM, Parmenter CA;
XX PI WPI; 1997-525714/48.
XX DR P-PSDB; AAM35742.
XX PT Evaluating risk of cervical dysplasia or cervical cancer - by detecting
variant form of human papilloma virus 16.
XX PS Claim 1; Col 15-24; 33pp; English.
XX CC Methods have been developed for distinguishing a subset of human
papilloma virus (HPV) that is associated with an increased risk of

```

CC developing cervical dysplasia or cervical cancer. The methods involve:
 CC (1) preparing a cervical sample to expose any HPV-16 E6 gene in the
 CC sample and determining if the base at position 350 of the E6 gene (see
 CC AAT94723 and AAT94724 for comparison) is T or G, where the presence of G
 CC at position 350 is associated with an increased risk of developing
 CC cervical dysplasia or cervical cancer; and (2) preparing a cervical
 CC sample to expose any HPV-16 E6 protein in the sample and determining if
 CC the amino acid at position 83 of the protein (see position 90 in AAW35741
 CC and AAW35742 for comparison) is Val or Leu, where the presence of Val at
 CC position 83 that is associated with an increased risk of developing
 CC cervical dysplasia or cervical cancer. The present sequence represents
 CC the variant nucleotide sequence for HPV-16 E6. The 350G variant
 CC correlates well with Pap scores: 350T:350G ratios among 45 HPV16 samples
 CC were 10:4 for negative Pap scores; 4:2 for CIN I, 1:6 for CIN II; 2:9 for
 CC CIN III; 0:3 for cancer. (Updated on 25-MAR-2003 to correct PF field.)
 XX
 SQ Sequence 7904 BP; 2601 A; 1377 C; 1510 G; 2416 T; 0 U; 0 Other;
 Query Match 100.0%; Score 28; DB 2; Length 7904;
 Best Local Similarity 100.0%; Pred. No. 0.047;
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 GAACGCAATACAAACCGTTGTGTG 28
 Db 368 GAACGCAATACAAACCGTTGTGTG 395
 RESULT 49
 AAT94723
 ID AAT94723 standard; DNA; 7904 BP.
 XX AAT94723;
 AC
 XX 25-MAR-2003 (revised)
 DT 16-FEB-1998 (first entry)
 DT
 XX
 DE Human papillomavirus type 16 E6 complete genome.
 XX
 XX Complete genome; circular; human papillomavirus type 16; HPV16 E6;
 KW cervical dysplasia; cervical cancer; cervical smear; ds.
 XX
 OS Human papillomavirus type 16.
 XX
 FH Key Location/Qualifiers
 FT TATA_signal 17..23
 FT TATA_signal 65..71
 FT TATA_signal 83..559
 FT CDS
 FT /*tag= c
 FT /product= "E6 transforming protein"
 FT /note= "E6 ORF from 65 to 559; putative"
 FT 350
 FT mutation
 FT /*tag= d
 FT /note= "Mutated from the normal T (as in this case) to G
 FT in the variant"
 FT 562..858
 FT CDS
 FT /*tag= e
 FT /product= "E7 transforming protein"
 FT /note= "E7 ORF 544 to 858; putative"
 FT 865..2813
 FT CDS
 FT /*tag= f
 FT /product= "E1 replication protein"
 FT /note= "E1 interrupted ORF from 859 to 2813; putative"
 FT 2755..3852
 FT CDS
 FT /*tag= g
 FT /product= "E2 regulatory protein"
 FT /note= "E2 ORF from 2725 to 3852; putative"
 FT 3332..3619
 FT CDS
 FT /*tag= h
 FT /product= "E4"
 FT /note= "E4 ORF from 3332 to 3619; putative"
 FT 3863..4099
 FT CDS

FT /*tag= i
 FT /product= "E5"
 FT /note= "E5 ORF from 3863 to 4099; putative"
 FT 4213..4218
 FT polyA_signal
 FT
 FT CDS
 FT /*tag= j
 FT /note= "Putative"
 FT 4235..5656
 FT /*tag= k
 FT /product= "L2 minor capsid protein"
 FT /note= "L2 ORF from 4133 to 5656; putative"
 FT 4289..4295
 FT TATA_signal
 FT /*tag= l
 FT 5559..7154
 FT CDS
 FT /*tag= m
 FT /product= "L1 major capsid protein"
 FT /note= "L1 ORF from 5526 to 7154; putative"
 FT 7260..7265
 FT polyA_signal
 FT /*tag= n
 FT
 FT US679509-A.
 PN
 XX 21-OCT-1997.
 XX
 XX 30-SEP-1994; 94US-00316239.
 PF
 XX 28-SEP-1993; 93US-00127906.
 XX
 XX (VINE-) UNIV NEW MEXICO STATE.
 PA
 XX Wheeler CM, Parmenter CA;
 PI
 XX WPI: 1997-525714/48.
 XX
 DR P-PSDB; AAW35741.
 DR
 PT Evaluating risk of cervical dysplasia or cervical cancer - by detecting
 PT variant form of human papilloma virus 16.
 XX
 XX Claim 1; Col 9-16; 33pp; English.
 PS
 XX Methods have been developed for distinguishing a subset of human
 CC papilloma virus (HPV) that is associated with an increased risk of
 CC developing cervical dysplasia or cervical cancer. The methods involve:
 CC (1) preparing a cervical sample to expose any HPV-16 E6 gene in the
 CC sample and determining if the base at position 350 of the E6 gene (see
 CC AAT94723 and AAT94724 for comparison) is T or G, where the presence of G
 CC at position 350 is associated with an increased risk of developing
 CC cervical dysplasia or cervical cancer; and (2) preparing a cervical
 CC sample to expose any HPV-16 E6 protein in the sample and determining if
 CC the amino acid at position 83 of the protein (see position 90 in AAW35741
 CC and AAW35742 for comparison) is Val or Leu, where the presence of Val at
 CC position 83 that is associated with an increased risk of developing
 CC cervical dysplasia or cervical cancer. The present sequence represents
 CC the reference nucleotide sequence for HPV-16 E6. The 350G variant
 CC correlates well with Pap scores: 350T:350G ratios among 45 HPV16 samples
 CC were 10:4 for negative Pap scores; 4:2 for CIN I, 1:6 for CIN II; 2:9 for
 CC CIN III; 0:3 for cancer. (Updated on 25-MAR-2003 to correct PF field.)
 CC
 XX
 SQ Sequence 7904 BP; 2601 A; 1377 C; 1509 G; 2417 T; 0 U; 0 Other;
 Query Match 100.0%; Score 28; DB 2; Length 7904;
 Best Local Similarity 100.0%; Pred. No. 0.047;
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 GAACGCAATACAAACCGTTGTGTG 28
 Db 368 GAACGCAATACAAACCGTTGTGTG 395
 RESULT 50
 AAX33881
 ID AAX33881 standard; DNA; 7904 BP.
 XX
 AC AAX33881;

```
XX 17-OCT-2003 (revised)
DT 25-JUN-1999 (first entry)
XX
XX HPV-16 genomic sequence.
DE
XX HPV-16; inhibitor; antisense oligonucleotide; E6/E7 gene; human;
KM keratinocyte; cervical cell; cervical tumour; PCR primer; beta-actin; ss.
XX
OS Human papillomavirus type 16.
XX
XX WO913071-A1.
XX
XX 18-MAR-1999.
XX
XX 03-SEP-1998; 98WO-US018320.
XX
XX 05-SEP-1997; 97US-00929140.
XX
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
XX Dipacolo J, Alvarez-Salas L;
XX
XX WPI, 1999-243727/20.
XX
XX New antisense oligonucleotide analogs for inhibiting growth of cervical
XX tumors.
XX
XX Disclosure; Page 28-32; 40pp; English.
XX
XX This sequence represents a sequence from the HPV-16 genome. This
XX invention relates to antisense oligonucleotide analogs (ONs) that have a
XX sequence complementary to a sequence of nucleotides 415-445 of human
XX papilloma virus-16 (HPV-16). The antisense ONs can be used to inhibit
XX expression of HPV gene E6/E7 in living cells, preferably human
XX keratinocytes or human cervical cells. They bind to E6/E7 mRNA in the
XX cell, prevent mRNA translation and promote mRNA degradation by
XX intracellular RNase H. They can be used for preventing transformation of
XX living cells by HPV. The antisense ONs are used particularly for
XX inhibiting the growth of cervical tumours. (Updated on 17-OCT-2003 to
XX standardise OS field)
XX
XX Sequence 7904 BP; 2601 A; 1377 C; 1509 G; 2417 T; 0 U; 0 Other;
SQ
XX
XX Query Match 100.0%; Score 28; DB 2; Length 7904;
XX Best Local Similarity 100.0%; Pred. No. 0.047;
XX Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GAACAGCAATACAAACCAACCGTTGTGTG 28
DB 368 GAACAGCAATACAAACCAACCGTTGTGTG 395
```

Search completed: May 24, 2006, 05:56:28
Job time : 335.037 secs

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 24, 2006, 05:47:01 ; Search time 1932.77 Seconds
(without alignments)
810.102 Million cell updates/sec

Title: US-10-601-913-5
Perfect score: 28
Sequence: 1 GAACAGCAATACAAACAACCGTTGTGTG 28

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 48236798 seqs, 27959665780 residues

Total number of hits satisfying chosen parameters: 96473596

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Listing first 1000 summaries

Database :

EST: *
1: gb_est1: *
2: gb_est3: *
3: gb_est4: *
4: gb_est5: *
5: gb_est6: *
6: gb_hic: *
7: gb_est2: *
8: gb_est7: *
9: gb_est8: *
10: gb_est9: *
11: gb_gsa1: *
12: gb_gsa2: *
13: gb_gsa3: *
14: gb_gsa4: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	28	100.0	244	3	B0106475
2	21.8	77.9	212	12	B2680296 PUBCT90TD
3	21.2	75.7	677	14	DE210433 Branchios
4	21.2	75.7	953	14	CNS03AIR
5	20.6	73.6	527	11	BZ321278
6	20.6	73.6	532	11	AZ412297
7	20.6	73.6	600	11	BZ314774
8	20.6	73.6	1058	11	CC274364
9	20	71.4	207	11	AQ075030
10	20	71.4	593	11	AQ265336
11	20	71.4	675	11	AQ505344
12	20	71.4	762	11	DR982710
13	20	71.4	846	14	CT237387
14	20	71.4	897	12	BG776822
15	20	71.4	1531	7	BE306182
16	19.8	70.7	502	2	BS512658
17	19.8	70.7	631	3	BP707553
18	19.8	70.7	637	3	BP700663
19	19.8	70.7	811	8	CO382721

20	19.6	70.0	311	8	CN471390
21	19.6	70.0	355	4	CB500937
22	19.6	70.0	487	10	DV085458
23	19.6	70.0	496	7	BE111725
24	19.6	70.0	579	11	AQ325004
25	19.6	70.0	658	8	CV232583
26	19.6	70.0	677	13	CL544775
27	19.6	70.0	682	5	CF234787
28	19.6	70.0	691	13	CZ749731
29	19.6	70.0	689	3	BM085832
30	19.6	70.0	735	3	BM087205
31	19.6	70.0	737	10	DM577165
32	19.6	70.0	753	4	CA064299
33	19.6	70.0	764	4	CA056575
34	19.6	70.0	782	4	CB643162
35	19.6	70.0	792	3	BU255194
36	19.6	70.0	793	3	BM087205
37	19.6	70.0	822	10	DM577166
38	19.6	70.0	832	12	CG219751
39	19.6	70.0	839	12	CG452337
40	19.6	70.0	959	12	CG219765
41	19.4	69.3	595	13	CZ104695
42	19.2	68.6	532	14	DE039251
43	19.2	68.6	524	3	BQ626720
44	19.2	68.6	524	3	BQ627018
45	19.2	68.6	534	3	BQ627107
46	19.2	68.6	573	3	BQ626527
47	19.2	68.6	707	3	BU220286
48	19.2	68.6	783	9	DR154444
49	19.2	68.6	817	3	BI153495
50	19.2	68.6	818	14	CT474369
51	19.2	68.6	934	11	B08236
52	19	67.9	219	7	AM113996
53	19	67.9	221	8	CV389718
54	19	67.9	270	7	BB327029
55	19	67.9	285	1	AA162702
56	19	67.9	302	1	AV227436
57	19	67.9	314	10	DV945298
58	19	67.9	380	7	BB793946
59	19	67.9	381	4	CT3040
60	19	67.9	389	4	BY675868
61	19	67.9	396	4	BY494189
62	19	67.9	397	4	BY507838
63	19	67.9	408	5	CJ271945
64	19	67.9	421	7	AV764126
65	19	67.9	422	7	AV763342
66	19	67.9	431	5	CJ301791
67	19	67.9	442	7	BB788790
68	19	67.9	453	2	BF856030
69	19	67.9	463	12	BZ659366
70	19	67.9	473	14	DE208132
71	19	67.9	480	12	BZ659378
72	19	67.9	495	11	AZ042156
73	19	67.9	499	11	BZ321269
74	19	67.9	500	12	BZ640370
75	19	67.9	506	12	BZ640364
76	19	67.9	536	14	DE008251
77	19	67.9	535	14	DE013352
78	19	67.9	549	11	AZ798253
79	19	67.9	550	9	DA013199
80	19	67.9	551	4	BM758434
81	19	67.9	554	2	BJ054696
82	19	67.9	558	10	DR390654
83	19	67.9	561	11	AZ390020
84	19	67.9	577	4	CB601885
85	19	67.9	598	4	BM739094
86	19	67.9	600	4	BM913349
87	19	67.9	603	3	BU258655
88	19	67.9	603	10	DV742883
89	19	67.9	610	2	BU079920
90	19	67.9	614	11	AZ443642
91	19	67.9	615	10	DV738589
92	19	67.9	627	4	CB527458

CN471390	h	Ab	Bra
CB500937	seal	pitho	
DV085458	337	-384	-1
BE111725	UI-R	Btl-	
AQ325004	mxg00	020B	
CV232583	MS0198	B2	
CL544775	OB	Ba007	
CF234787	PtAJXT001		
CZ749731	OC	Ba009	
BM085832	BM085832		
BM089764	BM089764		
DM577165	EST	seal	
CA064299	seal	rgb53	
CA056575	seal	rgb53	
CB643162	OSUNED03L		
BU255194	60374	6126	
BM087205	BM087205		
DM577166	EST	seal	
CG219751	OGMLQ26TH		
CG452337	OGVHQ26TV		
CG219765	OGMLQ26TV		
CZ104695	OM	Ba011	
DE039251	Branchios		
BQ626720	p226b03.y		
BQ627018	p230b06.y		
BQ627107	p231d03.y		
BQ626527	p211d01.y		
BU220286	603107	182	
DR154444	491103	65	
BI153495	602915	583	
CT474369	Sus	scrof	
B08236	P2020	SP6	
AM113996	MC617	mou	
CV389718	QV2	CT010	
BB327029	BB327029		
AA162702	mt42a02.x		
AV227436	AV227436		
DV945298	SB0300	02A1	
BB793946	BB793946		
CT3040	CT3040	Rice	
BY675868	BY675868		
BY494189	BY494189		
BY507838	BY507838		
CJ271945	CJ271945		
AV764126	AV764126		
AV763342	AV763342		
CJ301791	CJ301791		
BB788790	BB788790		
BF856030	CM4	FN019	
BZ659366	OCBP56TC		
DE208132	Branchios		
BZ659378	OCBP56TC		
AZ042156	RPCT	-23-3	
BZ321269	h21E05.y		
BZ640370	OCGAZ28TC		
BZ640364	OCGAZ28TC		
DE008251	Branchios		
DE013352	Branchios		
AZ798253	2M005	C02	
DA013199	DA013199		
BM758434	BM758434		
BJ054696	BJ054696		
DR390654	USDA	RP.1	
AZ390020	IM015H12		
CB601885	AGENCOURT		
BM739094	BM739094		
BM913349	BM913349		
BZ258655	603413	1723	
DV742883	ID0AFRBE		
BU079920	B1097920		
AZ443642	IM0238P03		
DV738589	ID0AFRBE		
CB527458	UI-M	FPY0-	

93	19	67.9	630	10	DV744567	DV744567	ID0A0G14A	C 166	18.6	66.4	629	2	BJ659994	BJ659994	Bu659994	Bu659994
C 94	19	67.9	632	2	BJ093863	BJ093863	BJ093863	C 167	18.6	66.4	636	11	BMBAC09K1	BMBAC09K1	Bu659994	Bu659994
C 95	19	67.9	633	2	DR077388	DR077388	LC_Cd3CF	C 168	18.6	66.4	644	2	BJ660441	BJ660441	Bu660441	Bu660441
C 96	19	67.9	639	12	CE145691	CE145691	CIgr-gr8-	C 169	18.6	66.4	648	14	BJ120743	BJ120743	BJ120743	BJ120743
97	19	67.9	642	3	BP224177	BP224177	BP224177	C 170	18.6	66.4	649	12	CL010096	CL010096	CL010096	CL010096
98	19	67.9	650	13	CM032566	CM032566	CM032566	C 171	18.6	66.4	652	8	CX294563	CX294563	CX294563	CX294563
99	19	67.9	701	3	BP224378	BP224378	BP224378	C 172	18.6	66.4	672	13	CZ789688	CZ789688	CZ789688	CZ789688
100	19	67.9	708	11	A2637338	A2637338	IM0496112	C 173	18.6	66.4	689	13	CZ733700	CZ733700	CZ733700	CZ733700
101	19	67.9	730	7	AM611425	AM611425	un67G04.x	C 174	18.6	66.4	701	12	CG789206	CG789206	CG789206	CG789206
C 102	19	67.9	731	8	CV559307	CV559307	UI-M-H20-	C 175	18.6	66.4	705	11	A2340332	A2340332	A2340332	A2340332
C 103	19	67.9	746	13	CT454665	CT454665	Sus BCroF	C 176	18.6	66.4	724	3	BM932548	BM932548	BM932548	BM932548
C 104	19	67.9	757	13	CZ161344	CZ161344	OR_Aba000	C 177	18.6	66.4	728	11	BM205750	BM205750	BM205750	BM205750
105	19	67.9	759	14	AG481510	AG481510	Mus muscu	C 178	18.6	66.4	740	12	CC747467	CC747467	CC747467	CC747467
C 106	19	67.9	766	7	BE034698	BE034698	ML03B11.M	C 179	18.6	66.4	762	3	BU247293	BU247293	BU247293	BU247293
C 107	19	67.9	778	12	CG857991	CG857991	ZMMBB026	C 180	18.6	66.4	770	12	CG397581	CG397581	CG397581	CG397581
C 108	19	67.9	785	5	CJ461017	CJ461017	CJ461017	C 181	18.6	66.4	778	12	CG833923	CG833923	CG833923	CG833923
C 109	19	67.9	796	10	DV624608	DV624608	94175.1.C	C 182	18.6	66.4	780	14	BM137533	BM137533	BM137533	BM137533
C 110	19	67.9	804	10	DR942184	DR942184	EST113372	C 183	18.6	66.4	780	14	CT138485	CT138485	CT138485	CT138485
111	19	67.9	809	1	AU121422	AU121422	AU121422	C 184	18.6	66.4	783	14	DU540408	DU540408	DU540408	DU540408
C 112	19	67.9	811	10	DM639962	DM639962	CUJ371-C2	C 185	18.6	66.4	790	12	CC901842	CC901842	CC901842	CC901842
C 113	19	67.9	831	10	DT631623	DT631623	DA572060	C 186	18.6	66.4	791	14	DUB37419	DUB37419	DUB37419	DUB37419
C 114	19	67.9	836	9	DA572060	DA572060	DA572060	C 187	18.6	66.4	794	12	CC140238	CC140238	CC140238	CC140238
C 115	19	67.9	845	11	BM086823	BM086823	RPCI-24-3	C 188	18.6	66.4	795	13	CZ773607	CZ773607	CZ773607	CZ773607
C 116	19	67.9	848	13	CZ540043	CZ540043	SRAA-aac12	C 189	18.6	66.4	801	13	DU241667	DU241667	DU241667	DU241667
C 117	19	67.9	861	10	DR690125	DR690125	SR108021	C 190	18.6	66.4	810	12	CC389170	CC389170	CC389170	CC389170
C 118	19	67.9	867	14	BM6010YK	BM6010YK	AL154140	C 191	18.6	66.4	813	10	DY335185	DY335185	DY335185	DY335185
119	19	67.9	868	7	BF218721	BF218721	601884948	C 192	18.6	66.4	821	2	BG622779	BG622779	BG622779	BG622779
C 120	19	67.9	902	7	BF699318	BF699318	602125881	C 193	18.6	66.4	825	14	CT430329	CT430329	CT430329	CT430329
C 121	19	67.9	908	14	CM082HRQ	CM082HRQ	AL197567	C 194	18.6	66.4	828	12	BZ288305	BZ288305	BZ288305	BZ288305
C 122	19	67.9	909	14	AG060538	AG060538	Oryza sat	C 195	18.6	66.4	834	13	CZ203693	CZ203693	CZ203693	CZ203693
C 123	19	67.9	929	13	CZ520536	CZ520536	GM216-63J1	C 196	18.6	66.4	836	10	CG428414	CG428414	CG428414	CG428414
C 124	19	67.9	999	12	CL037952	CL037952	CT0216-44H	C 197	18.6	66.4	840	10	DY334673	DY334673	DY334673	DY334673
C 125	19	67.9	1055	3	BU423914	BU423914	603203458	C 198	18.6	66.4	842	12	CC446754	CC446754	CC446754	CC446754
C 126	19	67.9	1175	9	DN695255	DN695255	CGX98-H12	C 199	18.6	66.4	868	8	CN761174	CN761174	CN761174	CN761174
C 127	19	67.9	1200	6	AK152391	AK152391	Mus muscu	C 200	18.6	66.4	913	4	BK447805	BK447805	BK447805	BK447805
C 128	19	67.9	1316	6	AK014470	AK014470	Mus muscu	C 201	18.6	66.4	945	5	CF542936	CF542936	CF542936	CF542936
C 129	19	67.9	1437	12	CC288945	CC288945	CH261-170	C 202	18.6	66.4	957	3	CZ316859	CZ316859	CZ316859	CZ316859
130	18.8	67.1	414	8	CO296980	CO296980	EK215606.	C 203	18.6	66.4	1012	12	CC223366	CC223366	CC223366	CC223366
131	18.8	67.1	755	12	BZ740194	BZ740194	OGFAY677M	C 204	18.6	66.4	1194	5	CD664329	CD664329	CD664329	CD664329
C 132	18.8	67.1	787	12	BZ740185	BZ740185	OGFAY677M	C 205	18.6	66.4	1541	14	AG568343	AG568343	AG568343	AG568343
C 133	18.8	67.1	789	13	CZ365568	CZ365568	ZMMBP0146	C 206	18.4	65.7	110	4	BM565158	BM565158	BM565158	BM565158
C 134	18.8	67.1	819	12	BZ288638	BZ288638	PUFHNA67B	C 207	18.4	65.7	163	4	BM535734	BM535734	BM535734	BM535734
135	18.8	67.1	881	14	DU589319	DU589319	CO_Ba008	C 208	18.4	65.7	186	12	CC565320	CC565320	CC565320	CC565320
136	18.8	67.1	934	12	BZ288640	BZ288640	PUFHN467D	C 209	18.4	65.7	230	4	BM583942	BM583942	BM583942	BM583942
C 137	18.8	67.1	1649	7	BF129403	BF129403	601811403	C 210	18.4	65.7	272	12	CE614632	CE614632	CE614632	CE614632
C 138	18.6	66.4	41	11	BZ286994	BZ286994	SALK_0230	C 211	18.4	65.7	365	5	CD190373	CD190373	CD190373	CD190373
C 139	18.6	66.4	270	13	CZ697407	CZ697407	OC_Ba001	C 212	18.4	65.7	381	7	BM168542	BM168542	BM168542	BM168542
C 140	18.6	66.4	276	8	CO049690	CO049690	Le_mxo_03	C 213	18.4	65.7	385	10	DM534437	DM534437	DM534437	DM534437
C 141	18.6	66.4	276	8	CO050928	CO050928	Le_mxo_21	C 214	18.4	65.7	386	1	AL039060	AL039060	AL039060	AL039060
142	18.6	66.4	342	13	CN575705	CN575705	rc57E07.x	C 215	18.4	65.7	393	11	AQ796835	AQ796835	AQ796835	AQ796835
143	18.6	66.4	342	13	CZ813758	CZ813758	OC_Ba018	C 216	18.4	65.7	398	9	DN830567	DN830567	DN830567	DN830567
C 144	18.6	66.4	366	8	CN576354	CN576354	rc63B01.x	C 217	18.4	65.7	411	11	AQ979872	AQ979872	AQ979872	AQ979872
C 145	18.6	66.4	389	3	BM952303	BM952303	rc57E07.Y	C 218	18.4	65.7	432	4	CD091062	CD091062	CD091062	CD091062
C 146	18.6	66.4	401	3	BQ563145	BQ563145	at05-d01	C 219	18.4	65.7	447	4	BM592737	BM592737	BM592737	BM592737
147	18.6	66.4	425	11	BI14896	BI14896	343F13.TP.C	C 220	18.4	65.7	448	4	BM533326	BM533326	BM533326	BM533326
148	18.6	66.4	448	11	AQ009738	AQ009738	CIT-HSP-2	C 221	18.4	65.7	454	8	CX280527	CX280527	CX280527	CX280527
149	18.6	66.4	474	11	AZ125836	AZ125836	OSUNB009	C 222	18.4	65.7	458	4	BM594086	BM594086	BM594086	BM594086
C 150	18.6	66.4	477	1	AA232260	AA232260	rx23c07.r	C 223	18.4	65.7	461	7	AM296554	AM296554	AM296554	AM296554
C 151	18.6	66.4	480	4	BM739507	BM739507	EX739507	C 224	18.4	65.7	470	2	BG993187	BG993187	BG993187	BG993187
152	18.6	66.4	485	1	AI230487	AI230487	EST227182	C 225	18.4	65.7	477	4	BM564260	BM564260	BM564260	BM564260
153	18.6	66.4	497	1	AI231705	AI231705	EST228393	C 226	18.4	65.7	481	8	CX257213	CX257213	CX257213	CX257213
C 154	18.6	66.4	510	14	CR967916	CR967916	mtb4-35D2	C 227	18.4	65.7	483	10	DR791285	DR791285	DR791285	DR791285
C 155	18.6	66.4	516	11	AQ877366	AQ877366	HS_2145.B	C 228	18.4	65.7	483	10	DM554073	DM554073	DM554073	DM554073
C 156	18.6	66.4	517	9	DR077652	DR077652	LC_Cd3CF	C 229	18.4	65.7	484	2	BJ696518	BJ696518	BJ696518	BJ696518
157	18.6	66.4	527	11	BS9650	BS9650	CIT-HSP-3J3	C 230	18.4	65.7	490	4	BM590610	BM590610	BM590610	BM590610
C 158	18.6	66.4	539	11	BH004287	BH004287	BMBAC07L1	C 231	18.4	65.7	497	4	BM577382	BM577382	BM577382	BM577382
C 159	18.6	66.4	571	1	AJ640761	AJ640761	AJ640761	C 232	18.4	65.7	502	2	BI349732	BI349732	BI349732	BI349732
C 160	18.6	66.4	581	3	BP354296	BP354296	BP354296	C 233	18.4	65.7	507	14	CR204334	CR204334	CR204334	CR204334
161	18.6	66.4	596	13	CZ841049	CZ841049	OC_Ba022	C 234	18.4	65.7	517	4	BK704401	BK704401	BK704401	BK704401
162	18.6	66.4	609	14	DUE21116	DUE21116	OC_Ba012	C 235	18.4	65.7	521	10	DR584434	DR584434	DR584434	DR584434
163	18.6	66.4	617	12	CC785359	CC785359	ZMMBB015	C 236	18.4	65.7	521	10	DR584434	DR584434	DR584434	DR584434
164	18.6	66.4	621	13	CZ708874	CZ708874	OC_Ba002	C 237	18.4	65.7	521	10	DR584434	DR584434		

239	18.4	65.7	528	4	BW534781	C 312	18.4	65.7	937	13	CZ521096
240	18.4	65.7	537	7	AW519969	313	18.4	65.7	941	12	CG265046
241	18.4	65.7	537	12	CG706619	C 314	18.4	65.7	946	10	DV006903
242	18.4	65.7	543	8	CV904193	C 315	18.4	65.7	955	12	CG394452
243	18.4	65.7	547	9	DA623668	C 316	18.4	65.7	975	12	CL069002
244	18.4	65.7	561	3	BW776568	C 317	18.4	65.7	980	13	CL286073
245	18.4	65.7	565	11	AZ027838	C 318	18.4	65.7	998	12	CG265055
246	18.4	65.7	567	9	DB036175	C 319	18.4	65.7	1103	5	CD498334
247	18.4	65.7	580	8	CX269212	C 320	18.4	65.7	1145	3	BU412541
248	18.4	65.7	584	9	DA126049	C 321	18.4	65.7	1197	14	AG186993
249	18.4	65.7	586	1	AI981187	C 322	18.4	65.7	1204	5	CD509495
250	18.4	65.7	590	2	B1649228	C 323	18.4	65.7	1290	3	BU368834
251	18.4	65.7	594	5	CF010859	C 324	18.2	65.0	147	4	CA188229
252	18.4	65.7	594	5	BX079501	C 325	18.2	65.0	204	4	CA074169
253	18.4	65.7	602	14	DE012579	C 326	18.2	65.0	279	10	R61674
254	18.4	65.7	604	10	DV535042	C 327	18.2	65.0	298	10	DM421348
255	18.4	65.7	611	4	BX079500	C 328	18.2	65.0	303	7	BE004518
256	18.4	65.7	614	10	DV952385	C 329	18.2	65.0	308	10	DM416756
257	18.4	65.7	617	11	AQ326024	C 330	18.2	65.0	338	7	BF001045
258	18.4	65.7	619	8	CV904964	C 331	18.2	65.0	344	2	BI043599
259	18.4	65.7	625	11	AZ026979	C 332	18.2	65.0	349	10	Z43070
260	18.4	65.7	626	11	DU840262	C 333	18.2	65.0	355	1	AV556319
261	18.4	65.7	629	4	CA055536	C 334	18.2	65.0	360	7	AM608740
262	18.4	65.7	638	9	CX888379	C 335	18.2	65.0	372	7	BF431728
263	18.4	65.7	646	1	AA603112	C 336	18.2	65.0	396	1	AA654845
264	18.4	65.7	646	7	BF166748	C 337	18.2	65.0	399	5	CF137006
265	18.4	65.7	649	5	CF428288	C 338	18.2	65.0	404	13	DU381977
266	18.4	65.7	654	13	CM303239	C 339	18.2	65.0	405	7	AV813536
267	18.4	65.7	656	14	BX242869	C 340	18.2	65.0	410	3	BP603580
268	18.4	65.7	658	4	EX085074	C 341	18.2	65.0	410	10	Z26572
269	18.4	65.7	667	8	CO837876	C 342	18.2	65.0	417	1	AA620667
270	18.4	65.7	667	8	CV938264	C 343	18.2	65.0	417	14	AY403684
271	18.4	65.7	679	2	B1523121	C 344	18.2	65.0	418	7	AV820481
272	18.4	65.7	700	8	CM090920	C 345	18.2	65.0	419	1	AA808883
273	18.4	65.7	703	2	BM171666	C 346	18.2	65.0	422	3	BP781429
274	18.4	65.7	711	4	BX085075	C 347	18.2	65.0	428	7	BF592278
275	18.4	65.7	712	4	BX082117	C 348	18.2	65.0	446	3	BP593514
276	18.4	65.7	718	2	BG427101	C 349	18.2	65.0	456	14	DE027250
277	18.4	65.7	721	8	CV908041	C 350	18.2	65.0	460	3	BP616531
278	18.4	65.7	722	14	AG359874	C 351	18.2	65.0	463	2	BI094462
279	18.4	65.7	731	8	CV937877	C 352	18.2	65.0	472	2	BJ455317
280	18.4	65.7	740	14	AG359827	C 353	18.2	65.0	473	1	AA859708
281	18.4	65.7	744	10	DR816150	C 354	18.2	65.0	483	2	BJ445566
282	18.4	65.7	746	12	CG932511	C 355	18.2	65.0	484	1	AA129646
283	18.4	65.7	749	8	CR538443	C 356	18.2	65.0	487	4	CA268829
284	18.4	65.7	752	14	AG313882	C 357	18.2	65.0	489	4	CB066954
285	18.4	65.7	764	5	CF222548	C 358	18.2	65.0	491	1	AA899200
286	18.4	65.7	767	9	CX806004	C 359	18.2	65.0	494	1	AA900526
287	18.4	65.7	773	14	BX174233	C 360	18.2	65.0	495	4	CB142360
288	18.4	65.7	777	9	CX799941	C 361	18.2	65.0	495	14	DX282918
289	18.4	65.7	783	12	CG273503	C 362	18.2	65.0	503	9	DN196565
290	18.4	65.7	786	14	AG543969	C 363	18.2	65.0	506	7	BE166435
291	18.4	65.7	786	6	AY111141	C 364	18.2	65.0	512	1	AL120965
292	18.4	65.7	794	2	B1904852	C 365	18.2	65.0	522	14	AY403683
293	18.4	65.7	795	7	BF120786	C 366	18.2	65.0	525	9	DB010538
294	18.4	65.7	796	14	CT089231	C 367	18.2	65.0	528	8	CV919403
295	18.4	65.7	800	7	BF167927	C 368	18.2	65.0	535	8	CR929777
296	18.4	65.7	801	4	BX073902	C 369	18.2	65.0	537	1	AA651004
297	18.4	65.7	804	11	B2388120	C 370	18.2	65.0	538	9	DA212267
298	18.4	65.7	809	9	DN532841	C 371	18.2	65.0	539	9	DB272081
299	18.4	65.7	811	9	CX803274	C 372	18.2	65.0	541	4	CA233257
300	18.4	65.7	822	9	DN531444	C 373	18.2	65.0	548	9	DB062794
301	18.4	65.7	847	12	CG365086	C 374	18.2	65.0	550	3	BP292526
302	18.4	65.7	854	9	DN044805	C 375	18.2	65.0	551	9	DB055760
303	18.4	65.7	854	13	DUI143834	C 376	18.2	65.0	552	3	BP220670
304	18.4	65.7	856	10	DT507970	C 377	18.2	65.0	556	9	DA811400
305	18.4	65.7	866	9	DN046363	C 378	18.2	65.0	559	9	DB182319
306	18.4	65.7	866	12	CG058723	C 379	18.2	65.0	561	9	DA563674
307	18.4	65.7	875	14	DUI40694	C 380	18.2	65.0	561	9	DA582329
308	18.4	65.7	880	9	DN111495	C 381	18.2	65.0	565	9	DA406339
309	18.4	65.7	881	13	CZ771074	C 382	18.2	65.0	566	9	DA430322
310	18.4	65.7	888	12	CG058726	C 383	18.2	65.0	568	9	DB222940
311	18.4	65.7	932	2	BI411242	C 384	18.2	65.0	573	9	DA326772

C 385	18.2	65.0	573	9	DA645571	DA645571	C 458	18.2	65.0	923	3	BO439332	BO439332
C 386	18.2	65.0	577	13	CM439888	CM439888	C 459	18.2	65.0	931	5	CF272246	CF272246
C 387	18.2	65.0	580	9	DA805032	DA805032	C 460	18.2	65.0	935	8	CO744483	CO744483
C 388	18.2	65.0	580	9	DB235482	DB235482	C 461	18.2	65.0	939	12	BZ796528	BZ796528
C 389	18.2	65.0	583	3	BP219199	BP219199	C 462	18.2	65.0	943	4	CA272360	CA272360
C 390	18.2	65.0	587	13	CM173235	CM173235	C 463	18.2	65.0	966	2	BG177968	BG177968
C 391	18.2	65.0	590	3	BU952508	BU952508	C 464	18.2	65.0	977	1	AL538186	AL538186
C 392	18.2	65.0	590	4	CA140756	CA140756	C 465	18.2	65.0	978	13	CZ838147	CZ838147
C 393	18.2	65.0	591	4	DA235634	DA235634	C 466	18.2	65.0	1002	4	BX328168	BX328168
C 394	18.2	65.0	594	10	DR376669	DR376669	C 467	18.2	65.0	1003	14	CNS07CSB	CNS07CSB
C 395	18.2	65.0	599	2	BG522609	BG522609	C 468	18.2	65.0	1043	12	CL117657	CL117657
C 396	18.2	65.0	608	1	AU005947	AU005947	C 469	18.2	65.0	1045	14	CNS01UBE	CNS01UBE
C 397	18.2	65.0	608	8	CO961294	CO961294	C 470	18.2	65.0	1054	3	BM800749	BM800749
C 398	18.2	65.0	616	3	BM341304	BM341304	C 471	18.2	65.0	1135	9	DN738343	DN738343
C 399	18.2	65.0	618	5	CD464626	CD464626	C 472	18.2	65.0	1139	5	CD498135	CD498135
C 400	18.2	65.0	632	7	BE835265	BE835265	C 473	18.2	65.0	1145	9	DN683027	DN683027
C 401	18.2	65.0	635	14	DE021735	DE021735	C 474	18.2	65.0	1208	3	BU962144	BU962144
C 402	18.2	65.0	638	9	CX528852	CX528852	C 475	18.2	65.0	1228	3	BU274693	BU274693
C 403	18.2	65.0	640	3	BP004973	BP004973	C 476	18.2	65.0	1283	10	DT958400	DT958400
C 404	18.2	65.0	640	4	BX508802	BX508802	C 477	18.2	65.0	1633	6	CNS0ACU7	CNS0ACU7
C 405	18.2	65.0	640	4	CA195268	CA195268	C 478	18.2	65.0	1665	6	CNS0ADU7	CNS0ADU7
C 406	18.2	65.0	643	10	DR373358	DR373358	C 479	18.2	65.0	2033	7	BE964846	BE964846
C 407	18.2	65.0	652	3	BU635676	BU635676	C 480	18.2	65.0	2350	6	CR602427	CR602427
C 408	18.2	65.0	654	10	DV083520	DV083520	C 481	18.2	65.0	3487	6	CR860117	CR860117
C 409	18.2	65.0	658	2	BI092246	BI092246	C 482	18.2	65.0	4781	6	CR749458	CR749458
C 410	18.2	65.0	666	8	CR771758	CR771758	C 483	18.2	65.0	4781	6	CR749458	CR749458
C 411	18.2	65.0	671	5	CK591123	CK591123	C 484	18.2	65.0	133	6	AK201353	AK201353
C 412	18.2	65.0	672	5	CD641207	CD641207	C 485	18.2	65.0	165	8	CR738298	CR738298
C 413	18.2	65.0	676	8	CO961265	CO961265	C 486	18.2	65.0	181	13	CL272036	CL272036
C 414	18.2	65.0	676	14	DX271231	DX271231	C 487	18.2	65.0	205	1	AV309083	AV309083
C 415	18.2	65.0	683	3	BM478345	BM478345	C 488	18.2	65.0	208	14	CT021072	CT021072
C 416	18.2	65.0	690	8	CA423284	CA423284	C 489	18.2	65.0	218	11	BL178047	BL178047
C 417	18.2	65.0	694	10	DT132404	DT132404	C 490	18.2	65.0	225	7	BM048018	BM048018
C 418	18.2	65.0	700	2	BI255430	BI255430	C 491	18.2	65.0	228	1	AL396409	AL396409
C 419	18.2	65.0	710	10	DM282542	DM282542	C 492	18.2	65.0	230	1	AV369045	AV369045
C 420	18.2	65.0	711	10	DM271030	DM271030	C 493	18.2	65.0	239	1	AV364657	AV364657
C 421	18.2	65.0	713	13	CZ774847	CZ774847	C 494	18.2	65.0	259	12	CG422372	CG422372
C 422	18.2	65.0	720	1	DM267359	DM267359	C 495	18.2	65.0	276	7	BB117256	BB117256
C 423	18.2	65.0	723	3	BG565570	BG565570	C 496	18.2	65.0	280	7	BB089624	BB089624
C 424	18.2	65.0	727	2	BG565570	BG565570	C 497	18.2	65.0	283	1	AV138118	AV138118
C 425	18.2	65.0	737	3	BM031052	BM031052	C 498	18.2	65.0	283	1	AV165915	AV165915
C 426	18.2	65.0	741	5	CD520101	CD520101	C 499	18.2	65.0	283	7	BB423881	BB423881
C 427	18.2	65.0	761	12	CG252895	CG252895	C 500	18.2	65.0	286	11	AQ096554	AQ096554
C 428	18.2	65.0	767	2	BG252571	BG252571	C 501	18.2	65.0	289	1	AV133243	AV133243
C 429	18.2	65.0	768	4	CA139649	CA139649	C 502	18.2	65.0	290	7	BB113403	BB113403
C 430	18.2	65.0	779	12	CG428568	CG428568	C 503	18.2	65.0	294	1	AA752403	AA752403
C 431	18.2	65.0	780	5	CD522765	CD522765	C 504	18.2	65.0	294	10	DM100231	DM100231
C 432	18.2	65.0	788	3	BF9642395	BF9642395	C 505	18.2	65.0	295	1	AV139966	AV139966
C 433	18.2	65.0	795	2	BF796991	BF796991	C 506	18.2	65.0	301	1	AV305877	AV305877
C 434	18.2	65.0	796	12	BZ738253	BZ738253	C 507	18.2	65.0	301	7	BB463309	BB463309
C 435	18.2	65.0	820	14	CNS03HOP	CNS03HOP	C 508	18.2	65.0	303	1	AV145461	AV145461
C 436	18.2	65.0	825	14	DU981676	DU981676	C 509	18.2	65.0	310	7	BB101150	BB101150
C 437	18.2	65.0	831	10	DT191750	DT191750	C 510	18.2	65.0	312	10	DM590603	DM590603
C 438	18.2	65.0	832	10	DT181174	DT181174	C 511	18.2	65.0	314	7	BB459285	BB459285
C 439	18.2	65.0	834	13	DU235681	DU235681	C 512	18.2	65.0	315	4	BY518836	BY518836
C 440	18.2	65.0	834	10	DR761505	DR761505	C 513	18.2	65.0	319	7	BB138243	BB138243
C 441	18.2	65.0	836	13	DU427808	DU427808	C 514	18.2	65.0	319	7	BB462386	BB462386
C 442	18.2	65.0	842	3	BU208078	BU208078	C 515	18.2	65.0	326	4	BY518900	BY518900
C 443	18.2	65.0	848	1	AU118142	AU118142	C 516	18.2	65.0	326	4	CA130357	CA130357
C 444	18.2	65.0	851	4	CA268870	CA268870	C 517	18.2	65.0	336	10	DM172543	DM172543
C 445	18.2	65.0	861	4	BQ431402	BQ431402	C 518	18.2	65.0	341	1	AI555397	AI555397
C 446	18.2	65.0	866	14	CR186150	CR186150	C 519	18.2	65.0	342	1	AA963352	AA963352
C 447	18.2	65.0	866	12	CG252903	CG252903	C 520	18.2	65.0	346	4	BM779292	BM779292
C 448	18.2	65.0	882	2	BG282937	BG282937	C 521	18.2	65.0	347	9	DN798724	DN798724
C 449	18.2	65.0	886	10	DM146986	DM146986	C 522	18.2	65.0	348	10	TO65780	TO65780
C 450	18.2	65.0	906	14	AG891226	AG891226	C 523	18.2	65.0	350	9	CX625815	CX625815
C 451	18.2	65.0	907	3	BQ425943	BQ425943	C 524	18.2	65.0	354	1	AM083571	AM083571
C 452	18.2	65.0	909	14	CNS03NCV	CNS03NCV	C 525	18.2	65.0	359	1	AL890379	AL890379
C 453	18.2	65.0	915	2	BG460620	BG460620	C 526	18.2	65.0	365	3	BU934530	BU934530
C 454	18.2	65.0	915	7	BE880238	BE880238	C 527	18.2	65.0	368	1	AI115603	AI115603
C 455	18.2	65.0	917	2	AL556802	AL556802	C 528	18.2	65.0	368	12	CE378300	CE378300
C 456	18.2	65.0	917	2	BM457887	BM457887	C 529	18.2	65.0	381	3	BM779578	BM779578
C 457	18.2	65.0	921	5	CF272506	CF272506	C 530	18.2	65.0	381	14	DU618221	DU618221

C 531	18	64.3	389	4	BY670564	BY670564	BY670564	604	18	64.3	587	8	CN947941	CN947941 020805AVB
C 532	18	64.3	393	4	BY420719	BY420719	BY420719	605	18	64.3	590	3	BU776576	BU776576 SUBCOD09
C 533	18	64.3	398	7	BB677135	BB677135	BB677135	606	18	64.3	593	2	BM177412	BM177412 aa181a09
C 534	18	64.3	398	7	BB766435	BB766435	BB766435	607	18	64.3	593	8	CN220637	CN220637 RAB067B05
C 535	18	64.3	399	11	AQ580525	AQ580525	RFC1-11-4	608	18	64.3	593	14	FR0039902	FR0039902
C 536	18	64.3	401	5	CF034372	CF034372	QCF4d12.Y	609	18	64.3	594	10	DV718070	DV718070 RVL8652 W
C 537	18	64.3	401	7	BB799914	BB799914	BB799914	610	18	64.3	595	1	AJ392478	AJ392478
C 538	18	64.3	402	4	BY477482	BY477482	BY477482	611	18	64.3	597	10	DM591820	DM591820 emu81-012
C 539	18	64.3	402	5	CF034665	CF034665	QCF8b06.Y	612	18	64.3	597	13	CZ249397	CZ249397 AIAA-aac9
C 540	18	64.3	402	10	DM591227	DM591227	emu81-008	613	18	64.3	598	3	BU485792	BU485792 603847534
C 541	18	64.3	404	1	AV541388	AV541388	AV541388	614	18	64.3	599	1	AL674692	AL674692
C 542	18	64.3	412	1	AA724860	AA724860	ah97c11.b	615	18	64.3	599	8	CK072448	CK072448 UCRCS08_2
C 543	18	64.3	417	10	DM591802	DM591802	emu81-012	616	18	64.3	599	8	CK072448	CK072448 UCRCS08_2
C 544	18	64.3	420	8	CV568148	CV568148	k425f05.Y	617	18	64.3	600	14	BX975708	BX975708 Reverse 8
C 545	18	64.3	427	3	BO555903	BO555903	HA036H07	618	18	64.3	600	14	CR099116	CR099116
C 546	18	64.3	433	1	AI447266	AI447266	mt01a11.x	619	18	64.3	601	3	BU337327	BU337327
C 547	18	64.3	433	1	DM091829	DM091829	CLPY3117.	620	18	64.3	601	8	CN938769	CN938769 010615AVB
C 548	18	64.3	444	5	CD240213	CD240213	ARELP51G5	621	18	64.3	602	8	CV852768	CV852768 gonad_EST
C 549	18	64.3	448	10	H40780	H40780	Y008b12.r1	622	18	64.3	603	4	CA054628	CA054628 aa11g52
C 550	18	64.3	454	2	BI804344	BI804344	H134C07 E	623	18	64.3	603	14	CITD10	AJ227432 Clona int
C 551	18	64.3	455	7	BF149049	BF149049	uv58b03.x	624	18	64.3	604	3	CA039615	CA039615 aa11nw01
C 552	18	64.3	459	1	AJ397662	AJ397662	AJ397662	625	18	64.3	605	3	BU398755	BU398755 603534057
C 553	18	64.3	463	9	CK625817	CK625817	GAN006C12	626	18	64.3	606	3	BU338275	BU338275 603514405
C 554	18	64.3	465	14	AC959426	AC959426	DY080PH11	627	18	64.3	606	3	BU352725	BU352725 603528626
C 555	18	64.3	471	4	CB130717	CB130717	K-EST0180	628	18	64.3	607	3	BU422137	BU422137 603232664
C 556	18	64.3	471	7	BF585741	BF585741	FM1_24_A0	629	18	64.3	607	9	DN855987	DN855987 4152699 B
C 557	18	64.3	476	5	CJ228478	CJ228478	CJ228478	630	18	64.3	611	3	BU463330	BU463330 603773864
C 558	18	64.3	477	4	BX554025	BX554025	BX554025	631	18	64.3	611	3	BU476643	BU476643 60384065
C 559	18	64.3	477	10	DT670077	DT670077	613dLT19A	632	18	64.3	611	9	CK909043	CK909043 JGI_BA01
C 560	18	64.3	491	4	CB637931	CB637931	OSUNEa05H	633	18	64.3	611	13	CZ429845	BU399113 603535384
C 561	18	64.3	494	4	CA347944	CA347944	679150 NC	634	18	64.3	612	3	BU477540	BU477540 603845676
C 562	18	64.3	497	7	BB772364	BB772364	BB772364	635	18	64.3	612	3	BU477540	BU477540 603845676
C 563	18	64.3	498	11	AZ366281	AZ366281	IM0115M14	636	18	64.3	613	13	CM880040	CM880040 Bn2457-2
C 564	18	64.3	505	2	BM509744	BM509744	1993d12.Y	637	18	64.3	613	13	BU339323	BU339323 603515635
C 565	18	64.3	506	2	BM117184	BM117184	L0847D06-	638	18	64.3	614	3	BU410489	BU410489 603158768
C 566	18	64.3	508	4	BM890089	BM890089	BM890089	639	18	64.3	614	9	CK734109	CK734109 1a08d09.Y
C 567	18	64.3	510	11	AQ328189	AQ328189	nbxb0042J	640	18	64.3	614	9	CK734109	CK734109 603372921
C 568	18	64.3	512	7	BE627511	BE627511	uu52e04.Y	641	18	64.3	615	4	CA382684	CA382684 662544 NC
C 569	18	64.3	513	2	BG051759	BG051759	FM1_58_A0	642	18	64.3	616	3	BU473834	BU473834 603365557
C 570	18	64.3	513	2	BG051759	BG051759	SGP154579	643	18	64.3	616	3	BU473834	BU473834 603365557
C 571	18	64.3	525	12	CG183504	CG183504	PJUFPL86TD	644	18	64.3	617	11	BH270376	BH270376 Cn230-139
C 572	18	64.3	528	12	CG183504	CG183504	PJUFPL86TD	645	18	64.3	617	3	BU421669	BU421669 603231785
C 573	18	64.3	532	14	AG951437	AG951437	DY080PH11	646	18	64.3	617	3	BU483961	BU483961 603847643
C 574	18	64.3	536	4	BM845887	BM845887	BM845887	647	18	64.3	621	7	AM173933	AM173933 f137d08.Y
C 575	18	64.3	537	2	BU654969	BU654969	BU654969	648	18	64.3	623	3	BU466995	BU466995 603373751
C 576	18	64.3	537	4	BY477927	BY477927	BY477927	649	18	64.3	624	3	BU475866	BU475866 603472339
C 577	18	64.3	541	8	CN511028	CN511028	AGENCOURT	650	18	64.3	625	3	BU479441	BU479441 603847631
C 578	18	64.3	542	4	BY475492	BY475492	BY475492	651	18	64.3	626	4	CA098635	CA098635 SCWCCL606
C 579	18	64.3	543	7	BB762562	BB762562	BB762562	652	18	64.3	627	3	BU350442	BU350442 603528992
C 580	18	64.3	546	3	BU457536	BU457536	60377566	653	18	64.3	627	3	BU419033	BU419033 603955670
C 581	18	64.3	547	11	AQ611675	AQ611675	HS_5087_B	654	18	64.3	628	4	CB270838	CB270838 Pgm1C.pko
C 582	18	64.3	548	5	CK330463	CK330463	H8219C10-	655	18	64.3	630	13	CL539356	CL539356 OB_Ba005
C 583	18	64.3	552	10	AQ426841	AQ426841	CITB1-EI-1	656	18	64.3	631	8	CV872926	CV872926 PDUt8a1108
C 584	18	64.3	553	11	DR909004	DR909004	USDA-FP_1	657	18	64.3	631	4	CA038691	CA038691 aa11nw00
C 585	18	64.3	556	3	BO558672	BO558672	H4053F05-	658	18	64.3	634	11	A0783407	A0783407 HS_5563_A
C 586	18	64.3	559	2	BM527247	BM527247	aa151E09.	659	18	64.3	641	2	BI763897	BI763897 603049663
C 587	18	64.3	559	4	BX265504	BX265504	BX265504	660	18	64.3	641	4	BM733064	BM733064 BM733064
C 588	18	64.3	563	7	BB906788	BB906788	BB906788	661	18	64.3	641	7	BE225590	BE225590 MD049 Me
C 589	18	64.3	564	7	BB909647	BB909647	BB909647	662	18	64.3	643	4	CB511309	CB511309 aa11nw50
C 590	18	64.3	567	14	BX965570	BX965570	Forward 8	663	18	64.3	646	3	BU309147	BU309147 603533808
C 591	18	64.3	567	14	DR37112S	DR37112S	ALX45570	664	18	64.3	648	1	AI230750	AI230750 EST227445
C 592	18	64.3	568	9	DN508079	DN508079	HL020060	665	18	64.3	649	1	AV304964	AV304964 AV304964
C 593	18	64.3	568	11	BH087067	BH087067	RPC1-24-3	666	18	64.3	649	13	CZ041955	CZ041955 OM_Ba002
C 594	18	64.3	570	4	BX087235	BX087235	EX087235	667	18	64.3	650	3	BU308795	BU308795 603538430
C 595	18	64.3	571	12	CE602308	CE602308	c1gr-g88-	668	18	64.3	651	4	CA053862	CA053862 aa11g5b3
C 596	18	64.3	572	3	BU345640	BU345640	604173089	669	18	64.3	651	14	DE002403	DE002403 Branch108
C 597	18	64.3	574	8	CK258231	CK258231	1311396 N	670	18	64.3	652	4	BX087234	BX087234 BX087234
C 598	18	64.3	575	7	BB905012	BB905012	BB905012	671	18	64.3	658	1	AJ455332	AJ455332 AJ455332
C 599	18	64.3	576	3	BU766095	BU766095	STRAN01	672	18	64.3	658	7	BF693429	BF693429 602081085
C 600	18	64.3	585	11	AQ836855	AQ836855	HS_3463_A	673	18	64.3	660	13	CZ048822	CZ048822 OM_Ba003
C 601	18	64.3	585	13	DJ396589	DJ396589	109842121	674	18	64.3	662	8	CV719843	CV719843 UCRCS08_0
C 602	18	64.3	586	3	BP125395	BP125395	BP125395	675	18	64.3	663	11	AZ941346	AZ941346 2M0201L05
C 603	18	64.3	586	12	BZ868329	BZ868329	CH240_237	676	18	64.3	664	12	CG114957	CG114957 PUFO842TB

677	18	64.3	666	3	BU478335	750	18	64.3	780	7	BF239645	BF239645	601906892
678	18	64.3	666	11	BH505127	751	18	64.3	781	11	BZ471312	BZ471312	BONOC0707F
C 679	18	64.3	667	9	CX845727	752	18	64.3	781	11	CG834123	CG834123	ZMMBC014
680	18	64.3	668	13	CZ173806	753	18	64.3	782	10	DV725494	DV725494	RYL17674
C 681	18	64.3	668	13	CN505GNO	754	18	64.3	782	10	BU118416	BU118416	603142667
682	18	64.3	672	4	CA506869	755	18	64.3	783	13	CZ568449	CZ568449	CO21ALU-F
683	18	64.3	676	3	BU360503	756	18	64.3	785	9	CX913086	CX913086	JGI_CANA3
C 684	18	64.3	676	14	AG175996	757	18	64.3	785	12	CG397273	CG397273	ZMMBC001
C 685	18	64.3	678	14	CB511412	758	18	64.3	786	3	BU773192	BU773192	SJBFIC05
686	18	64.3	678	14	CT246510	759	18	64.3	787	9	CX399045	CX399045	CT463723
C 687	18	64.3	679	3	BU403409	760	18	64.3	787	12	CC004433	CC004433	PUD0227D
C 688	18	64.3	680	3	BM576558	761	18	64.3	793	3	BU389636	BU389636	603511550
C 689	18	64.3	680	7	BB326822	762	18	64.3	794	13	DU031157	DU031157	12408 TOM
C 690	18	64.3	680	9	CX360499	763	18	64.3	796	12	BZ990181	BZ990181	PUGGW727B
C 691	18	64.3	681	4	CA048121	764	18	64.3	797	13	CZ349316	CZ349316	ZMMBR0124
692	18	64.3	681	11	BZ465576	765	18	64.3	803	3	BU237755	BU237755	603401114
C 693	18	64.3	682	9	DN959160	766	18	64.3	803	8	CR413724	CR413724	CA413724
C 694	18	64.3	684	14	AG218826	767	18	64.3	803	12	CC697423	CC697423	OGVAL95TV
C 695	18	64.3	685	5	CF837284	768	18	64.3	806	11	BZ194850	BZ194850	CH230-370
C 696	18	64.3	685	8	CV719844	769	18	64.3	809	10	DV614902	DV614902	EST121789
C 697	18	64.3	687	8	CO140355	770	18	64.3	813	9	CX399046	CX399046	JGI_XZT35
C 698	18	64.3	689	13	CM378730	771	18	64.3	813	12	CG458197	CG458197	PUGFW73TD
699	18	64.3	691	3	BM977444	772	18	64.3	814	14	CT476371	CT476371	Sus scroF
700	18	64.3	694	1	AJ455323	773	18	64.3	818	9	CX961292	CX961292	JGI_CAO1
C 701	18	64.3	694	13	CZ197729	774	18	64.3	820	9	CX961291	CX961291	JGI_CAO1
702	18	64.3	695	13	DU591467	775	18	64.3	821	10	DT347307	DT347307	JGI_CAX1
703	18	64.3	696	7	BF468449	776	18	64.3	822	8	CX333400	CX333400	JGI_XZT69
C 704	18	64.3	699	12	BZ748780	777	18	64.3	827	3	BU110689	BU110689	603126603
C 705	18	64.3	700	14	AG148120	778	18	64.3	827	10	DT451076	DT451076	JGI_CABR9
706	18	64.3	702	8	CN383191	779	18	64.3	827	12	CG255521	CG255521	OG0BR36TV
707	18	64.3	704	11	AQ530921	780	18	64.3	830	11	BH432496	BH432496	BOGFQ40TR
C 708	18	64.3	713	11	AZ562166	781	18	64.3	831	14	CT259175	CT259175	Sus scroF
C 709	18	64.3	716	8	CB509606	782	18	64.3	832	14	BX223701	BX223701	Danto xar
710	18	64.3	716	8	CX031704	783	18	64.3	834	14	BX135201	BX135201	Danto xar
C 711	18	64.3	717	8	CX258232	784	18	64.3	835	12	CG455202	CG455202	PUGFW73TB
C 712	18	64.3	721	5	CK935576	785	18	64.3	838	14	CT434010	CT434010	Sus scroF
C 713	18	64.3	722	12	CG903456	786	18	64.3	841	2	BM049606	BM049606	603624678
714	18	64.3	723	12	BZ990184	787	18	64.3	844	2	DU599659	DU599659	OO_BA009
715	18	64.3	723	13	CZ660450	788	18	64.3	848	14	BF683842	BF683842	963074A12
C 716	18	64.3	728	3	BU465521	789	18	64.3	852	10	DM018775	DM018775	EST122773
C 717	18	64.3	733	5	CF833657	790	18	64.3	857	14	CT428450	CT428450	Sus scroF
718	18	64.3	733	11	BH897937	791	18	64.3	859	10	DR876755	DR876755	JGI_CAB17
719	18	64.3	733	14	DX289424	792	18	64.3	860	10	DT407137	DT407137	JGI_CAB18
C 720	18	64.3	734	3	BU437725	793	18	64.3	862	8	CR439116	CR439116	CR439116
C 721	18	64.3	735	9	CX890781	794	18	64.3	863	11	BH163643	BH163643	ENTOR07TF
722	18	64.3	739	13	CL598493	795	18	64.3	866	14	CT134853	CT134853	Sus scroF
C 723	18	64.3	740	12	CC085680	796	18	64.3	867	12	CT198458	CT198458	ZMMBC025
724	18	64.3	741	5	CF837283	797	18	64.3	867	14	DU599746	DU599746	OO_BA009
C 725	18	64.3	742	5	CF833676	798	18	64.3	867	14	CR053018	CR053018	Forward B
C 726	18	64.3	743	9	CX668747	799	18	64.3	870	9	DN910105	DN910105	Cuq4_2 CO
727	18	64.3	744	3	BO850013	800	18	64.3	873	8	CV504104	CV504104	70620_1 M
728	18	64.3	744	3	DN061462	801	18	64.3	874	14	DU629500	DU629500	OO_BA013
C 729	18	64.3	744	14	AG552560	802	18	64.3	876	12	CG122462	CG122462	PUGFW73TB
C 730	18	64.3	745	2	BU015509	803	18	64.3	877	9	DN935437	DN935437	AGBNCOURT
C 731	18	64.3	745	3	BU482146	804	18	64.3	878	3	BU119751	BU119751	603142358
C 732	18	64.3	745	14	SBC561077	805	18	64.3	879	7	BF679455	BF679455	602153258
733	18	64.3	746	9	CX668748	806	18	64.3	885	12	CC985642	CC985642	ZUABE91TV
C 734	18	64.3	747	14	DX300262	807	18	64.3	886	10	DM171047	DM171047	CLVZ2077_
C 735	18	64.3	748	3	BM968609	808	18	64.3	889	4	CB096250	CB096250	AP53-RPF
736	18	64.3	752	5	CF731569	809	18	64.3	889	14	CN801763	CN801763	ALU61616
C 737	18	64.3	753	3	BU382903	810	18	64.3	891	14	CR018070	CR018070	Forward B
738	18	64.3	757	11	AQ210780	811	18	64.3	894	13	BU398249	BU398249	603534238
C 739	18	64.3	760	5	CK937042	812	18	64.3	894	12	CG097102	CG097102	PUGFW73TD
C 740	18	64.3	763	13	CZ464506	813	18	64.3	894	13	CZ216526	CZ216526	AIAA-aeF2
C 741	18	64.3	765	9	CX783951	814	18	64.3	896	8	CO933136	CO933136	AGBNCOURT
742	18	64.3	768	2	BU604677	815	18	64.3	897	14	AG170095	AG170095	Pan troG1
743	18	64.3	768	12	CG422244	816	18	64.3	897	14	CR174287	CR174287	Forward B
C 744	18	64.3	770	11	AZ611475	817	18	64.3	901	14	AG834289	AG834289	Oryza sat
745	18	64.3	771	5	CF342419	818	18	64.3	903	2	BF964790	BF964790	602266038
C 746	18	64.3	771	5	CF342419	819	18	64.3	908	12	BZ786559	BZ786559	PUGFW40TD
747	18	64.3	774	4	CA918136	820	18	64.3	910	10	DV619521	DV619521	EST122251
C 748	18	64.3	775	5	CK358592	821	18	64.3	910	12	BZ786554	BZ786554	PUGFW40TB
C 749	18	64.3	778	12	CG114958	822	18	64.3	912	12	CG255510	CG255510	OG0BR36TH

823	18	64.3	916	8	CO761123	brain EST	896	17.8	63.6	634	13	CZ159101	OO_Ba000
824	18	64.3	921	3	BU119102	BU119102 603141831	897	17.8	63.6	635	4	BX677900	BX677900
825	18	64.3	921	10	DT803481	DT803481 126536955	898	17.8	63.6	637	13	CZ079135	CZ079135
826	18	64.3	924	4	BX759355	BX759355	899	17.8	63.6	639	11	AQ447489	AQ447489
827	18	64.3	930	10	DT799123	DT799123 126437153	900	17.8	63.6	640	8	CZ041507	CZ041507
828	18	64.3	931	12	CC187727	CC187727 CH261-145	901	17.8	63.6	649	13	CN586235	CN586235
829	18	64.3	936	12	CG103302	CG103302 PUJB611TB	902	17.8	63.6	649	14	DUS68542	DUS68542
830	18	64.3	943	12	CG276996	CG276996 CG3AM26TV	903	17.8	63.6	652	14	DUS59876	DUS59876
831	18	64.3	944	2	BI464720	BI464720 603202340	904	17.8	63.6	653	14	DUS51688	DUS51688
832	18	64.3	945	3	BU136737	BU136737 603125395	905	17.8	63.6	663	10	DM010486	DM010486
833	18	64.3	948	10	DM597296	DM597296 CGX129-GO	906	17.8	63.6	660	11	BZ012237	BZ012237
834	18	64.3	954	10	DT804783	DT804783 127377390	907	17.8	63.6	666	14	DUS67527	DUS67527
835	18	64.3	958	2	BG104489	BG104489 602311126	908	17.8	63.6	690	13	CZ610987	CZ610987
836	18	64.3	986	13	CM920306	CM920306 EDCARF4TF	909	17.8	63.6	692	13	CZ669460	CZ669460
837	18	64.3	986	10	CC294866	CC294866 CH261-89D	910	17.8	63.6	693	14	CT360222	CT360222
838	18	64.3	1045	12	DM037992	DM037992 CFW275-E0	911	17.8	63.6	697	14	DUS65679	DUS65679
839	18	64.3	1058	9	DN791375	DN791375 909365526	912	17.8	63.6	711	14	BU548774	BU548774
840	18	64.3	1062	9	DN654554	DN654554 CEC11-G12	913	17.8	63.6	712	14	BX136427	BX136427
841	18	64.3	1074	12	CC222198	CC222198 CH261-71C	914	17.8	63.6	716	4	CB454422	CB454422
842	18	64.3	1296	8	CV861758	CV861758 gonad_EST	915	17.8	63.6	719	12	CR326126	CR326126
843	18	64.3	1319	9	DN700141	DN700141 CLJ27-G08	916	17.8	63.6	720	5	CK962095	CK962095
844	18	64.3	1384	6	AY812526	AY812526 SchiBcObo	917	17.8	63.6	727	14	DUS628804	DUS628804
845	18	64.3	1411	14	AJ854675	AJ854675 BrabstCa	918	17.8	63.6	731	13	CZ060890	CZ060890
846	18	64.3	1422	3	BM911789	BM911789 AGENCOURT	919	17.8	63.6	735	14	DUS66629	DUS66629
847	18	64.3	1552	12	BZ566603	BZ566603 CH261-164	920	17.8	63.6	735	14	DUS74890	DUS74890
848	18	64.3	1560	12	CC211685	CC211685 CH261-186	921	17.8	63.6	737	14	DUS615328	DUS615328
849	18	64.3	2217	14	AG278863	AG278863 Mus muscu	922	17.8	63.6	756	14	DUS40537	DUS40537
850	18	64.3	2325	6	AK046630	AK046630 Mus muscu	923	17.8	63.6	758	14	DX261153	DX261153
851	18	64.3	2325	6	BC052999	BC052999 Mus muscu	924	17.8	63.6	759	5	CK459828	CK459828
852	18	64.3	3255	6	AK077691	AK077691 Mus muscu	925	17.8	63.6	764	14	DUS48461	DUS48461
853	18	64.3	3255	6	AA489292	AA489292 ab36f04.r	926	17.8	63.6	767	14	DUS48469	DUS48469
854	17.8	63.6	160	3	BP561315	BP561315	927	17.8	63.6	770	13	CZ881509	CZ881509
855	17.8	63.6	252	3	BP561315	BP561315	928	17.8	63.6	773	14	DUS70939	DUS70939
856	17.8	63.6	261	14	DUS48246	DUS48246 OO_Ba002	929	17.8	63.6	776	7	BF244931	BF244931
857	17.8	63.6	286	13	CL632137	CL632137 XY03425-5	930	17.8	63.6	785	8	CO724139	CO724139
858	17.8	63.6	317	10	DY127558	DY127558 001106BEM	931	17.8	63.6	801	10	DT835590	DT835590
859	17.8	63.6	320	14	DUS68279	DUS68279 OO_Ba007	932	17.8	63.6	818	14	DUS621056	DUS621056
860	17.8	63.6	330	14	DUS68279	DUS68279 OO_Ba007	933	17.8	63.6	824	14	DUS99884	DUS99884
861	17.8	63.6	333	13	CZ625803	CZ625803 OM_Ba016	934	17.8	63.6	832	10	DT840391	DT840391
862	17.8	63.6	346	13	CZ110313	CZ110313 OM_Ba012	935	17.8	63.6	834	14	DUS61363	DUS61363
863	17.8	63.6	394	13	CZ649764	CZ649764 OM_Ba009	936	17.8	63.6	842	4	BX838246	BX838246
864	17.8	63.6	399	13	CZ090952	CZ090952 OM_Ba009	937	17.8	63.6	844	10	DV908135	DV908135
865	17.8	63.6	402	14	DUS45877	DUS45877 OO_Ba002	938	17.8	63.6	885	10	DV927878	DV927878
866	17.8	63.6	408	11	AQ581856	AQ581856 RRC1-11.4	939	17.8	63.6	893	10	DUS64800	DUS64800
867	17.8	63.6	433	11	CZ491883	CZ491883 OM_BB013	940	17.8	63.6	902	8	CO880504	CO880504
868	17.8	63.6	439	1	AA313711	AA313711 EST185579	941	17.8	63.6	902	14	DUS99934	DUS99934
869	17.8	63.6	449	13	CZ601288	CZ601288 OM_Ba012	942	17.8	63.6	912	12	CC262538	CC262538
870	17.8	63.6	452	13	CZ047632	CZ047632 OM_Ba003	943	17.8	63.6	1067	10	DM032270	DM032270
871	17.8	63.6	459	1	AJ667832	AJ667832	944	17.8	63.6	1094	12	CC284139	CC284139
872	17.8	63.6	462	9	CV798041	CV798041 UGI_CAAJ1	945	17.8	63.6	1069	12	CC284139	CC284139
873	17.8	63.6	497	5	CD286676	CD286676 12_P8_abd	946	17.8	63.6	1112	12	CC268770	CC268770
874	17.8	63.6	505	13	CZ622708	CZ622708 OM_Ba016	947	17.8	63.6	1189	9	DN660749	DN660749
875	17.8	63.6	514	13	CZ643334	CZ643334 OM_Ba018	948	17.8	63.6	1191	9	DN663210	DN663210
876	17.8	63.6	526	13	CZ057765	CZ057765 OM_Ba005	949	17.8	63.6	1195	9	DN657552	DN657552
877	17.8	63.6	535	13	CZ613791	CZ613791 OM_Ba014	950	17.8	63.6	1211	9	DN702449	DN702449
878	17.8	63.6	536	14	DUS70392	DUS70392 OO_Ba005	951	17.8	63.6	1230	9	DN659521	DN659521
879	17.8	63.6	542	14	CR822673	CR822673 GR0AA52C	952	17.8	63.6	1271	5	CD503083	CD503083
880	17.8	63.6	548	13	CZ671177	CZ671177 OM_Ba022	953	17.8	63.6	92	11	AA253980	AA253980
881	17.8	63.6	550	13	CZ086633	CZ086633 OM_Ba008	954	17.8	63.6	100	7	BF478272	BF478272
882	17.8	63.6	554	14	DUS603020	DUS603020 OO_Ba010	955	17.6	62.9	110	11	AZ950433	AZ950433
883	17.8	63.6	554	14	DUS603020	DUS603020 OO_Ba010	956	17.6	62.9	126	7	AM238566	AM238566
884	17.8	63.6	559	5	CV2089780	CV2089780 OM_Ba009	957	17.6	62.9	134	11	BH772224	BH772224
885	17.8	63.6	562	5	CV792511	CV792511 885454 MA	958	17.6	62.9	149	5	CF623083	CF623083
886	17.8	63.6	562	8	CV869399	CV869399 PDUB1066	959	17.6	62.9	166	7	BF286942	BF286942
887	17.8	63.6	568	13	CZ624289	CZ624289 OM_Ba016	960	17.6	62.9	169	7	AM238351	AM238351
888	17.8	63.6	579	13	CZ096767	CZ096767 OM_Ba010	961	17.6	62.9	159	7	AM238351	AM238351
889	17.8	63.6	587	14	LBAR093D09	LBAR093D09	962	17.6	62.9	178	10	W14610	W14610
890	17.8	63.6	588	8	CV870158	CV870158 PDUB1075	963	17.6	62.9	187	7	AM238686	AM238686
891	17.8	63.6	601	3	BP700624	BP700624	964	17.6	62.9	199	7	AM238686	AM238686
892	17.8	63.6	602	7	BF045997	BF045997 BP250003B	965	17.6	62.9	201	2	BG020487	BG020487
893	17.8	63.6	613	14	DUS52880	DUS52880 OO_Ba003	966	17.6	62.9	212	1	AI115395	AI115395
894	17.8	63.6	626	14	DUS37792	DUS37792 OO_Ba000	967	17.6	62.9	212	1	AI115395	AI115395
895	17.8	63.6	626	14	DUS37792	DUS37792 OO_Ba000	968	17.6	62.9	212	1	AI115395	AI115395

c 969	17.6	62.9	220	7	BF420392	BF420392 UI-R-BJ2-
c 970	17.6	62.9	230	7	AM919342	AM919342 EST50646
c 971	17.6	62.9	231	7	AW919340	AW919340 X245F08.X
c 972	17.6	62.9	235	1	AV150166	AV150166 AV150166
c 973	17.6	62.9	251	7	BE678370	BE678370 dF54N03.X
c 974	17.6	62.9	254	5	CF609345	CF609345 INF1001.0
c 975	17.6	62.9	259	13	CZ668582	CZ668582 OM_Ba022
c 976	17.6	62.9	272	5	CK401544	CK401544 PBUI6 Mo02
c 977	17.6	62.9	273	2	BJ743794	BJ743794 BJ743794
c 978	17.6	62.9	275	1	AL838713	AL838713 AL838713
c 979	17.6	62.9	277	7	BB244244	BB244244 BB244244
c 980	17.6	62.9	278	7	AW919651	AW919651 EST350955
c 981	17.6	62.9	286	1	AI104033	AI104033 EST213322
c 982	17.6	62.9	287	10	DM319969	DM319969 LRAGE0447
c 983	17.6	62.9	291	10	DM311174	DM311174 LRAGE0460
c 984	17.6	62.9	292	1	AV025034	AV025034 AV025034
c 985	17.6	62.9	294	7	BB659810	BB659810 GM7000108
c 986	17.6	62.9	303	7	BF287189	BF287189 EST451780
c 987	17.6	62.9	309	7	BB928602	BB928602 BB928602
c 988	17.6	62.9	311	13	CZ079176	CZ079176 OM_Ba007
c 989	17.6	62.9	314	3	BQ146350	BQ146350 NF047803P
c 990	17.6	62.9	316	10	DM294675	DM294675 LRAGE0575
c 991	17.6	62.9	322	10	DM297618	DM297618 LRAGE0604
c 992	17.6	62.9	322	12	CC178214	CC178214 KC289 Bay
c 993	17.6	62.9	326	10	DM300401	DM300401 LRAGE0622
c 994	17.6	62.9	326	13	CZ078596	CZ078596 OM_Ba007
c 995	17.6	62.9	328	1	AA958500	AA958500 ua13604.x
c 996	17.6	62.9	330	2	BI395675	BI395675 EST51546
c 997	17.6	62.9	333	8	CV676083	CV676083 mu44f04.K
c 998	17.6	62.9	333	1	AI703827	AI703827 UI-R-AC1
c 999	17.6	62.9	346	1	AA899801	AA899801 UI-R-B0-d
1000	17.6	62.9	360	4	CB377556	CB377556 CMaE1_38

ALIGNMENTS

RESULT 1
LOCUS BQ106475 244 bp mRNA linear EST 16-APR-2002
DEFINITION fc2102.e Rose Petals (Fragrant Cloud) Lambda Zap Express Library
VERSION BQ106475
KEYWORDS BQ106475.1 GI:20156137
SOURCE EST.
ORGANISM Rosa hybrid cultivar
Rosa hybrid cultivar
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids I; Rosales; Rosaceae; Rosoideae; Rosa.
1 (bases 1 to 244)
Gulerman,I., Shalit,M., Menda,N., Plestun,D., Dafny-Yellin,M.,
Shalev,G., Bar,E., Davydov,O., Ovadis,M., Emanuel,M., Wang,J.,
Adam,Z., Pichersky,E., Lewinsohn,E., Zamir,D., Vainstein,A. and
Weiser,D.
Rosa Scent: Genomics Approach to Discovering Novel Floral
Fragrance-Related Genes
Plant Cell 14 (10), 2325-2338 (2002)
Contact: Naama Menda
Petal Genomics
Faculty of Agricultural, Food and Environmental Quality Sciences,
The Hebrew University of Jerusalem
P.O. Box 12, Rehovot, 76100, Israel
Tel: 972 8 9489 683
Fax: 972 8 9468 265
Email: shaham@agri.huji.ac.il
Seq primer: T3 forward.
Location/Qualifiers
1..244
/organism="Rosa hybrid cultivar"
/mol_type="mRNA"
/strain="Fragrant Cloud"
/db_xref="taxon:128735"

FEATURES
source

FEATURES
source

/clone="fc2102.e"
/issue_type="Petals"
/dev_stage="Young open flower at stage four"
/clone_lib="Rose Petals (Fragrant Cloud) Lambda Zap
Express Library"
/note="Vector: pBKCMV; Site_1: EcoRI; Site_2: XhoI"

Qy 1 GAACGCAATACACAAACCGTTGTG 28
Db 49 GAACGCAATACACAAACCGTTGTG 76

RESULT 2
LOCUS BZ680296/c 212 bp DNA linear GSS 05-FEB-2003
DEFINITION PUBCT90TD ZM 0.6 1.0 KB Zea mays genomic clone ZMMBTA025P12,
Genomic survey sequence.
ACCESSION BZ680296
VERSION BZ680296.1 GI:28235036
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 212)
Whitelaw,C.A., Quackenbush,J., Van Aken,S., Uteback,T.,
Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
Bennetzen,J.
Maize Genomics Consortium
Unpublished (2003)
Contact: Cathy Whitelaw
TIGR
912 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TF
Class: sheared ends.
Location/Qualifiers
1..212
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone_lib="ZMMBTA025P12"
/clone_lib="ZM 0.6 1.0 KB"
/note="Vector: pCR4-TOPO; Site_1: EcoRI; 0.6-1.0 kb high
cot selected genomic DNA library"

FEATURES
source

ORIGIN
Query Match 77.9%; Score 21.8; DB 12; Length 212;
Best Local Similarity 92.0%; Pred. No. 99;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 AACGCAATACACAAACCGTTGTG 26
Db 122 ATCGCAATACACAAACCGTTGTG 98

RESULT 3
LOCUS DE210433/c 677 bp DNA linear GSS 19-AUG-2005
DEFINITION Branchiostoma floridae DNA, clone: CH302-100_A12_TJ, genomic survey
sequence.
ACCESSION DE210433
VERSION DE210433.1 GI:73597487
KEYWORDS GSS.

SOURCE Branchiostoma floridae (Florida lancelet)
ORGANISM Branchiostoma floridae
Bukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae;
Branchiostoma.

REFERENCE 1
AUTHORS Fujiiyama, A., Toyoda, A., Kuroki, Y. and Sakaki, Y.
TITLE BAC end sequences of CHORI-302 Amphioxus library
JOURNAL Published Only in Database (2005)
REFERENCE 2 (bases 1 to 677)
AUTHORS Fujiiyama, A.
TITLE Direct Submission
JOURNAL Submitted (12-AUG-2005) Ageo Fujiiyama, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC); 1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa, 230-0045, Japan (E-mail: fujiiyama@gsc.riken.jp, URL: http://etc.gsc.riken.jp/, Tel: 81-45-503-9111, Fax: 81-45-503-9170)
COMMENT PRIMERS
Sequencing : TV
LIBRARY
Vector : PTARBAC2.1
R.Site 1 : EcoRI
R.Site 2 : EcoRI.
Location/Qualifiers
1. .677
/organism="Branchiostoma floridae"
/mol_type="genomic DNA"
/db_xref="taxon:7739"
/clone="CH302-100_A12_TV"
/sex="male"
/tissue type="sperm"
/clone_lib="CHOR10302 Amphioxus genomic BAC library"
/note="Common name: amphioxus"

ORIGIN
Query Match 75.7%; Score 21.2; DB 14; Length 677;
Best Local Similarity 88.5%; Pred. No. 2e+02;
Matches 23; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AACGCAATACAAACCGTTGTGT 27
Db 170 AATAGCAATCAACAAACCGTTGTGT 145

RESULT 4
CNS03AIR 953 bp DNA linear GSS 01-SEP-2000
LOCUS Tetradon nigroviridis genome survey sequence T7 end of clone
DEFINITION 010C15 of library G from Tetradon nigroviridis, genomic survey sequence.
ACCESSION AL235260 GI:7894395
VERSION AL235260.1 GI:7894395
KEYWORDS GSS; genome survey sequence.
SOURCE Tetradon nigroviridis
ORGANISM Tetradon nigroviridis
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes; Tetraodontidae; Tetraodon.

REFERENCE 1
AUTHORS Roest Crolius, H., Jalllon, O., Dasilva, C., Bouneau, L., Fisher, C., Bernot, A., Fizesen, C., Wincker, P., Brottier, P., Quetier, F., Saurin, W. and Weissenbach, J.
TITLE Estimate of human gene number provided by genome-wide analysis using Tetradon nigroviridis DNA sequence
JOURNAL Nat. Genet. 25 (2), 235-238 (2000)
REFERENCE 2
AUTHORS Roest Crolius, H., Jalllon, O., Dasilva, C., Ozouf-Costaz, C., Fizesen, C., Fischer, C., Bouneau, L., Billaud, A., Quetier, F., Saurin, W., Bernot, A. and Weissenbach, J.
TITLE Characterization and repeat analysis of the compact genome of the freshwater pufferfish Tetradon nigroviridis
JOURNAL Genome Res. 10 (7), 939-949 (2000)

PUBMED 10899143
REFERENCE 3 (bases 1 to 953)
ORGANISM Genoscope.
TITLE Direct Submission
JOURNAL Submitted (12-APR-2000) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT This sequence is a single read and was generated as part of a large scale clone-end sequencing project of the Tetradon nigroviridis genome. For more information, please take a look at http://www.genoscope.cns.fr/tetraodon.

FEATURES
source
1. .953
/organism="Tetradon nigroviridis"
/mol_type="genomic DNA"
/db_xref="taxon:99883"
/clone="010C15"
/clone_lib="G"
/note="Genoscope sequence ID : COBG010AB08LP1 end : T7"

ORIGIN
Query Match 75.7%; Score 21.2; DB 14; Length 953;
Best Local Similarity 88.5%; Pred. No. 2.1e+02;
Matches 23; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AACGCAATACAAACCGTTGTGT 27
Db 18 AATGAAATACAAACCGTTGTGT 43

RESULT 5
BZ321278 527 bp DNA linear GSS 06-NOV-2002
LOCUS h221g05.g1 WGS-2may9f (JM107 adapted methyl filtered) Zea mays
DEFINITION genomic clone h221g05 5', genomic survey sequence.
ACCESSION BZ321278
VERSION BZ321278.1 GI:24697359
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoidae; Andropogoneae; Zea.
1 (bases 1 to 527)
Rabinowitz, P.D., O'Shaughnessy, A.L., Ballia, V., Dedhia, N., Katzenburger, F., King, L., Miller, B., Muller, S., Nascento, L., Zlatavarn, T., McCombie, W.R. and Martienssen, R.A.
Genomic shotgun sequences from Zea mays (methyl-filtered)
Unpublished (2002)
Contact: W. Richard McCombie
Lita Annenberg Hazen Genome Sequencing Center
Cold Spring Harbor Laboratory
PO Box 100, Cold Spring Harbor, NY 11724, USA
Tel: 516 367 8874
Fax: 516 367 8884
Email: mcombie@cshl.org
Plate: h221 row: 9 column: 05
Seq primer: -21M13UnivRev
Class: shotgun
High quality sequence stop: 527.
Location/Qualifiers
1. .527
/organism="Zea mays"
/mol_type="genomic DNA"
/cultivar="B73"
/db_xref="taxon:4577"
/clone="h221g05"
/lab_host="JM107 or DH5a"
/clone_lib="WGS-2may9f (JM107 adapted methyl filtered)"
/note="Organ: Immature ears; Site 1: Xba I; Site 2: Xba I; The vector was digested with XbaI and one nucleotide was added by fill in in the recessive 3' end. The genomic DNA

was nebulized, end repaired, adaptor ligated and size fractionated using sephadex. The resulting fragments were between 0.8 and 3 kb and were cloned into the vector (x/y reads in M13mp19, b/g reads in pUC19). The same ligation was transformed in either JM107 or DH5a.

ORIGIN

Query Match 73.6%; Score 20.6; DB 11; Length 527;
Best Local Similarity 85.2%; Pred. No. 3.6e+02;
Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 GAACAGCAATACAAACCGTTGTGT 27
|||||
Db 215 GAACACCAATTAACAAATCGTGTGT 241

RESULT 6
AZ412297/c 532 bp DNA linear GSS 03-OCT-2000
LOCUS 1M0185M08R Mouse 10kb plasmid UGCM1 library Mus musculus genomic
DEFINITION clone UGCM1M0185M08 R, genomic survey sequence.
ACCESSION AZ412297
VERSION AZ412297.1 GI:10536310
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Bukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Muridae; Murinae; Mus.
1 (bases 1 to 532)
Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C.,
Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T.,
Reilly, M., Rose, M., Rose, R., Stokes, R., Tingey, A., von
Niederhausern, A. and Wright, D., Weiss, R.
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
Unpublished (2000)
Contact: Robert B. Weiss
University of Utah Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: dunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0185 row: M column: 08
Seq primer: CACACAGAAACAGCTATGAC
Class: plasmid ends
High quality sequence stop: 532.
Location/Qualifiers
1. 532
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UGCM1M0185M08"
/sex="Male"
/lab_host="R. Coli strain XL10-Gold, T1-resistant, F-"
/clone_lib="Mouse 10kb plasmid UGCM1 library"
/note="Vector: PMD42nv; Purified genomic DNA from M.
laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adaptored DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of PMD42 (g14732114|gb|AF129072.1), a copy-number

inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptored mouse DNA was annealed to adaptored vector DNA, and transformed into chemically-competent E.coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

ORIGIN

Query Match 73.6%; Score 20.6; DB 11; Length 532;
Best Local Similarity 85.2%; Pred. No. 3.6e+02;
Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 AACAGCAATACAAACCGTTGTGTG 28
|||||
Db 515 AACAGCAATACAAACCGTTGTGTG 489

RESULT 7
BZ314774/c 600 bp DNA linear GSS 06-NOV-2002
LOCUS h221g05.b1 WGS-Zmay9f (JM107 adapted methyl filtered) Zea mays
DEFINITION genomic clone h221g05 5', genomic survey sequence.
ACCESSION BZ314774
VERSION BZ314774.1 GI:24683701
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 600)
Rabinowicz, P.D., O'Shaughnessy, A.L., Bailly, V., Dedhia, N.,
Katzenburger, F., King, L., Miller, B., Muller, S., Naciminto, L.,
Zutavern, T., McCombie, W.R. and Martienssen, R.A.
Genomic shotgun sequences from Zea mays (methyl-filtered)
Unpublished (2002)
Contact: W. Richard McCombie
Lita Annenberg Hazen Genome Sequencing Center
Cold Spring Harbor Laboratory
PO Box 100, Cold Spring Harbor, NY 11724, USA
Tel: 516 367 8884
Fax: 516 367 8874
Email: mccombie@cshl.org
Plate: h221 row: g column: 05
Seq primer: -21M13UnivFwd
Class: shotgun
High quality sequence stop: 600.
Location/Qualifiers
1. 600
/organism="Zea mays"
/mol_type="genomic DNA"
/cultivar="B73"
/db_xref="taxon:4577"
/clone="h221g05"
/lab_host="JM107 or DH5a"
/clone_lib="WGS-Zmay9f (JM107 adapted methyl filtered)"
/note="Organ: immature ears; Site 1: Xba I, Site 2: Xba I,
The vector was digested with XbaI and one nucleotide was
added by fill in in the recessive 3' end. The genomic DNA
was nebulized, end repaired, adaptor ligated and size
fractionated using sephadex. The resulting fragments were
between 0.8 and 3 kb and were cloned into the vector
(x/y reads in M13mp19, b/g reads in pUC19). The same
ligation was transformed in either JM107 or DH5a."

QY 1 GAACAGCAATACAAACCGTTGTGT 27
|||||
Db 381 GAACACCAATTAACAAATCGTGTGT 355

FEATURES

source

1. 532
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UGCM1M0185M08"
/sex="Male"
/lab_host="R. Coli strain XL10-Gold, T1-resistant, F-"
/clone_lib="Mouse 10kb plasmid UGCM1 library"
/note="Vector: PMD42nv; Purified genomic DNA from M.
laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adaptored DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of PMD42 (g14732114|gb|AF129072.1), a copy-number

ORIGIN

Query Match 73.6%; Score 20.6; DB 11; Length 600;
Best Local Similarity 85.2%; Pred. No. 3.6e+02;
Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 GAACAGCAATACAAACCGTTGTGT 27
|||||
Db 381 GAACACCAATTAACAAATCGTGTGT 355

RESULT 8
CC274364/c 1058 bp DNA 1linear GSS 13-MAY-2003
LOCUS CH261-20G8.SP6.1 CH261 Gallus gallus genomic clone CH261-20G8,
DEFINITION genomic survey sequence.
ACCESSION CC274364
VERSION CC274364.1 GI:30631425
KEYWORDS GSS.
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Phasiantidae; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
REFERENCE 1 (bases 1 to 1058)
AUTHORS Krenitzki, C., Higginbotham, J., Wylie, K., Carter, J., McPherson, J.,
Warren, W., Graves, T., Mardis, E. and Wilson, R.
TITLE Gallus gallus BAC End Reads
JOURNAL Unpublished (2003)
COMMENT Contact: Richard K. Wilson
Genome Sequencing Center
Washington University School of Medicine
Email: submissions@wustl.edu
Insert Length: 182000 Std Error: 0.00
Seq primer: SP6 ATTAGTGTACTATAG
Class: BAC ends
High quality sequence start: 26
High quality sequence stop: 571.
Location/Qualifiers
1..1058
/organism="Gallus gallus"
/mol_type="genomic DNA"
/strain="Red Jungle Fowl"
/db_xref="taxon:9031"
/clone="CH261-20G8"
/sex="female"
/cell_line="UCD001, inbred 256"
/clone_lib="CH261"
/note="Vector: pPARBAC2.1; Site_1: EcoRI; Site_2: EcoRI;
CH261 Female Chicken library - for library and clone
ordering information: <http://www.chori.org/bacpac>"

ORIGIN
Query Match 73.6%; Score 20.6; DB 12; Length 1058;
Best Local Similarity 85.2%; Pred. No. 3.8e+02;
Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 GAACAGCAATACCAACCGCTGTCTG 27
Db 422 GAACAGCAATACCAACCGCTGTCTAT 396

RESULT 9
AA075030/c 207 bp DNA 1linear GSS 20-AUG-1998
LOCUS CIT-HSP-2364011.TR CIT-HSP Homo sapiens genomic clone 2364011,
DEFINITION genomic survey sequence.
ACCESSION AA075030
VERSION AA075030.1 GI:3435601
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominae; Homo.
REFERENCE 1 (bases 1 to 207)
AUTHORS Adams, M.D., Rounsley, S.D., Zhao, S., Baas, S., Linher, K., Golden, K.,
Berry, K., Granger, D., Sun, E., Wible, C., Shizuya, H., Simon, M. and
Venter, J.C.
TITLE Use of a random human BAC End Sequence Database for Sequence-Ready
JOURNAL Map Building
Unpublished (1998)

COMMENT Other GSSs: CIT-HSP-2364011.TF
Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
Seq primer: M13 Reverse
Class: BAC ends.

FEATURES
source
1..207
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="2364011"
/sex="Male"
/cell_type="Sperm"
/clone_lib="CIT-HSP"
/note="Vector: pBelBAC11; Site_1: HindIII; Site_2:
HindIII"

ORIGIN
Query Match 71.4%; Score 20; DB 11; Length 207;
Best Local Similarity 82.1%; Pred. No. 6.1e+02;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 GAACCGAATACCAACCGCTGTGTG 28
Db 127 GAACCGAATACCAACCGCTGTGTG 100

RESULT 10
AQ265336 593 bp DNA 1linear GSS 27-OCT-1998
LOCUS CITBI-EI-2506L6.TF CITBI-EI Homo sapiens genomic clone 2506L6,
DEFINITION genomic survey sequence.
ACCESSION AQ265336
VERSION AQ265336.1 GI:3793536
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominae; Homo.
REFERENCE 1 (bases 1 to 593)
AUTHORS Adams, M.D., Rounsley, S.D., Zhao, S., Baas, S., Linher, K., Golden, K.,
Berry, K., Granger, D., Sun, E., Wible, C., Shizuya, H., Simon, M. and
Venter, J.C.
TITLE Use of a random human BAC End Sequence Database for Sequence-Ready
JOURNAL Map Building
Unpublished (1998)
COMMENT Other GSSs: CITBI-EI-2506L6.TF
Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
Seq primer: M13-21
Class: BAC ends.

FEATURES
source
1..593
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

```

/clone="2506L6"
/sex="male"
/cell_type="sperm"
/clone_id="CTR1-X1"
/notes="Vector: pBeloBAC11, Site_1: EcoRI, Site_2: EcoRI
Caltech Human BAC Library D"

```

```

Query Match      71.4%; Score 20; DB 11; Length 593;
Best Local Similarity 82.1%;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```

QY 1 GAACAGCATACACAAACCGTTGTG 28
Db 285 GTACAGCAATACCAACACAGGTGTG 312

RESULT	11
LOCUS	AO505344
DEFINITION	AO505344 675 bp DNA linear GSS 29-APR-1998 RPCT-11-281E13.TU RPCT-11 Homo sapiens genomic clone RPCT-11-281E13, genomic survey sequence.
ACCESSION	AO505344
VERSION	AO505344.1 GI:4710091
KEYWORDS	GSS.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens

REFERENCE
1 (bases 1 to 675)
AUTHORS
Zhao, S., Adams, M.D., Nierman, W., Malek, J., de Jong, P. and
Venter, J.C.

Map Building
Unpublished (1997)
Other_GSSs: RPCT-11-201B13_TV

COMMENT

Other GSs: RPCI-11-281B3, TV
Contact: Shaying Zhao, William Nierman, Mark Adams
Department of Bukaroyotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850
Tel: 301 838 0200
Fax: 301 838 0208
Email: hbeetlgr.org
Clones are derived from the human BAC library RPCI-11. For BAC
library availability, please contact Pieter de Jong
(pieterdejong.med.buffalo.edu). Clones may be purchased from
BACPC Resources (<http://bacpac.med.buffalo.edu/ordering>) or from
Research Genet cs (info@resgen.com). BAC end search page:
http://www.tigr.org/cdb/hungen/bac_end_search/bac_end_search.html
Seq primer: Sp6
Class: BAC ends.

```

FEATURES
    source
        Location/Qualifiers
            1..675
                /organism="Homo sapiens"
                /mol_type="genomic DNA"
                /db_xref="GDB:7607628"
                /db_xref="taxon:9606"
                /clone="RPCT-11-281E13"
                /sex="Male"
                /cell_type="Lymphocytes"
                /clone_11b="RPCT-11"
                /note="Vector: pBAC3.6; Site_1: EcoRI; Site_2: EcoRI
                RPCT11 Human Male BAC Library"

```

Query Match	71.4%	Score 20	DB 11	length 675
Best Local Similarity	82.1%	Pred. No.	6.7e+02	
Matches 23	Conservative 0	Mismatches 5	Indels 0	Gaps 0

```

QY      1 GAACAGCAATACAAACAAACCGTTGTGTG 28
          |||||
Db      349 GTACAGCAATACCAACACACAGGTGTGTG 376

```

RESULT	12
DR982710/c	
LOCUS	762 bp mRNA linear EST 03-AUG-2005
DEFINITION	JGI AOKF63 fwd AOKP Acropora palmata spawned eggs Acropora palmata
ACCESSION	DR982710
VERSION	DR982710.1 GI:71777416
KEYWORDS	Est.
SOURCE	Acropora palmata

REFERENCES
Eukaryota: Metazoa: Cnidaria; Anthozoa; Hexacorallia; Scleractinia;
Astrocoemilia; Actinoptidae; Actinoptera.
1 (bases 1 to 762)
Schwarz, J.A., Brokstein, P., Manohar, C., Coffroth, M.A., Szmidt, A.

TITLE	JOURNAL	COMMENT
Coral-Symbiodinium EST Project	Unpublished (2005)	Other_ESTs: JGI_AOKF633.rev

Contact: Schwarz, JA, Medina, M.
Evolutionary Genomics
DOE Joint Genome Institute
2800 Mitchell Drive, Walnut Creek, CA 94596, USA
Tel.: 925-296-5823
Email: j.schwarz@lbl.gov
cDNA Library Preparation: DOE Joint Genome Institute
<http://www.jgi.doe.gov>

DNA Sequencing: DOE Joint Genome Institute: <http://www.jgi.doe.gov>
 Naming Conventions: EST name is generated by the concatenation of the UGI Clone Id and the direction of sequencing. The suffix '.fwd' indicates a forward sequencing read of the insert. It does not necessarily reflect the orientation of the insert.
 Plate: AOPF 0005 row: a column: 16
 High quality sequence stop: 701.

```

FEATURES
    source
        location/Qualifiers
            1..762
                /organism="Acropora palmata"
                /mol_type="mRNA"
                /db_xref="taxon:6131"
                /clone="AOKF63"
                /dev_stage="Freshly spawned eggs"
                /lab_host="Electromax DH10B"
                /clone_1lb="AOKF Acropora palmata spawned eggs"
                /note="Vector: pDNR-Lib; Site_1: SfiI; Site_2: SfiI. The
                library was prepared from total RNA using the Creator
                SMART cDNA Library Construction Kit with the ID-PCR method
                to amplify the cDNA. Amplified cDNA was digested with
                SfiI, size selected for >400bp, and ligated into the
                pDNR-Lib vector."

```

Query Match	71.4%	Score 20;	DB 10;	Length 763;
Best Local Similarity	82.1%	Pred. No. 6.8e+02;		
Matches 23;	Conservative 0;	Mismatches 5;	Indels 0;	Gaps 0

OY 1 GAACAGCAATACAAACAACCGTTGTGTC 28
 |||||
Db 476 GAACAGCAATGCCTTAAACCATTTGTGTC 449
 |||||

RESULT	13
LOCUS	CT237387
DEFINITION	Sus scrofa genomic clone CH242-181F11, genomic survey sequence.
ACCESSION	CT237387
VERSION	CT237387.1
KEYWORDS	GSS.
SOURCE	Sus scrofa (pig)
ORGANISM	Sus scrofa

Bukaryota, Metazoa; Chordata, Craniata, Euteleostomi, Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Suina; Suidae, Sus.

REFERENCE 1 (bases 1 to 846)
 AUTHORS Humphray,S.J., Plumb,R.W. and Durham,J.L.
 TITLE Direct Submission
 JOURNAL Submitted (01-NOV-2005) The Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humphrey@sanger.ac.uk Unpublished

COMMENT This sequence was generated from the T7 end of BAC 181F11. 181F11 is part of the CHOBI-242 BAC library created by P. de Jong. Further details: http://www.sanger.ac.uk/Projects/S_scrofa/.
 Location/Qualifiers
 source 1..846
 /organism="Sus scrofa"
 /mol_type="genomic DNA"
 /db_xref="taxon:9823"
 /clone="CH242-181F11"
 /tissue_type="white blood cells"
 /note="vector pTARBAC1.3_BamHI sex female"

ORIGIN
 Query Match 71.4%; Score 20; DB 14; Length 846;
 Best Local Similarity 82.1%; Pred. No. 6.9e+02;
 Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 GAACAGCAATACAAACCGTTGTGTG 28
 Db 48 GAACGCAATACAAACCGTTGTGTG 75

RESULT 14
 LOCUS BG776822 897 bp mRNA linear EST 15-MAY-2001
 DEFINITION 602664002F1 NIH_MGC_59 Homo sapiens CDNA clone IMAGE:4809169 5', mRNA sequence.
 ACCESSION BG776822
 VERSION BG776822.1 GI:14047139
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 897)
 NIH-MGC <http://mgc.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgabs-remail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: CLONTECH Laboratories, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLML) DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLML at: <http://image.llnl.gov>
 Plate: LLM6541 row: d column: 02
 High quality sequence stop: 459.
 Location/Qualifiers
 source 1..897
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4809169"
 /tissue_type="mucoepidermoid carcinoma"
 /lab_host="DH10B (T1 phage-resistant)"
 /clone_1lb="NIH MGC 59"
 /note="Organ: lung; Vector: pDNR-LIB (Clontech); Site_1: SfiI (ggcgccctcgcc); Site_2: SfiI (ggccattatggc); 3' Double-stranded CDNA was prepared from cell line RNA. 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CAAGCGCAATTAAGGC-3' and 3' adaptor sequence: 5'-ATTCTAGAGCGCCAGCGCCACATG-dt(30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average

insert size 1.65 kb (range 0.9-4.0 kb). 15/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA). Note: this is a NIH_MGC Library."

ORIGIN
 Query Match 71.4%; Score 20; DB 2; Length 897;
 Best Local Similarity 82.1%; Pred. No. 6.9e+02;
 Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 GAACAGCAATACAAACCGTTGTGTG 28
 Db 527 GAGCAGCAGACGACGACCGTTGTGTG 554

RESULT 15
 LOCUS BE306182/c 1531 bp mRNA linear EST 26-OCT-2000
 DEFINITION 601101310F1 NCI_CGAP_Lu29 Mus musculus CDNA clone IMAGE:3493587 5', mRNA sequence.
 ACCESSION BE306182
 VERSION BE306182.1 GI:9159012
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Muridae; Murinae; Mus.
 1 (bases 1 to 1531)
 NIH-MGC <http://mgc.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgabs-remail.nih.gov
 Tissue Procurement: Gilbert Smith, Ph.D.
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLML) DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLML at: <http://image.llnl.gov>
 Plate: LLM6541 row: d column: 04
 High quality sequence stop: 101.
 Location/Qualifiers
 source 1..1531
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="Czech II"
 /db_xref="taxon:10090"
 /clone="IMAGE:3493587"
 /tissue_type="spontaneous tumor, metastatic to mammary. Stem cell origin."
 /lab_host="DH10B"
 /clone_1lb="NCI CGAP Lu29"
 /note="Organ: lung; Vector: pCMV-SPORT6; Site_1: SalI; Site_2: NotI; Cloned unidirectionally. Primer: Oligo dt. Library constructed by Life Technologies. Investigator providing samples: Gilbert Smith, NIH"

ORIGIN
 Query Match 71.4%; Score 20; DB 7; Length 1531;
 Best Local Similarity 82.1%; Pred. No. 7.2e+02;
 Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 GAACAGCAATACAAACCGTTGTGTG 28
 Db 157 GAGCAGCAGACGACGACCGTTGTGTG 130

RESULT 16
 LOCUS BG512658 502 bp mRNA linear EST 28-MAR-2001

DEFINITION dad31b04.x1 wellcome CRC PCS107 tropicallis St10-12 Xenopus
 tropicallis cDNA clone IMAGE:4440919 3' similar to SW:SET_HUMAN
 001105 SET PROTEIN ; mRNA sequence.
ACCESSION BG512658
VERSION BG512658.1 GI:13483315
KEYWORDS EST.
SOURCE Xenopus tropicallis (western clawed frog)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Xenopus tropicallis
REFERENCE 1 (bases 1 to 502)
 Clifton,S., Johnson,S.L., Blumberg,B., Song,J., Hillier,L.,
 Pape,D., Martin,J., Wylie,T., Underwood,K., Theising,B., Bowers,Y.,
 Peterson,B., Gibbons,M., Harvey,N., Ritter,E., Jackson,Y., McCann,R.,
 Waterston,R. and Wilson,R.
 Washu Xenopus EST project, 1999
 Unpublished (1999)
TITLE Other ESTs: dad31b04.y1
JOURNAL Contact: Sandy Clifton, Ph.D.
COMMENT Washu Xenopus EST project, 1999
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Library: est@watson.wustl.edu
 Library constructed by A. Zorn and J. Mason (Wellcome/CRC
 Institute). DNA Sequencing by: Washington University Genome
 Sequencing Center
 Clone distribution: Xenopus clones from this library are available
 through the I.M.A.G.E. Consortium/BLNI at: info@image.llnl.gov
 Possible reversed clone: similarity on wrong strand
 Seq primer: -40UP from Gibco
 High quality sequence stop: 486.
FEATURES Location/Qualifiers
 1..502
 /organism="Xenopus tropicallis"
 /mol_type="mRNA"
 /db_xref="taxon:8354"
 /clone="IMAGE:4440919"
 /issue_type="whole embryo, stages 10-12"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="Wellcome CRC PCS107 tropicallis St10-12"
 /note="Vector: PCS107; Site 1: NotI; Site 2: EcoRI; CDNA's
 were oligo-dT primed and directionally cloned. Average
 insert size 1.5 kb. range 0.5-4 kb. Library constructed by
 A. Zorn and J. Mason (Wellcome/CRC Institute)."
ORIGIN
 Query Match 70.7%; Score 19.8; DB 2; Length 502;
 Best Local Similarity 91.3%; Pred. No. 8.1e+02;
 Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 GAACGACATACACAAACCGTT 23
 |||||
 |||||
Db 250 GAACGAAATACACAAACCGTT 272
 |||||
 |||||
RESULT 17
LOCUS BP707553 631 bp mRNA linear EST 19-JUL-2004
DEFINITION BP707553 Osada Taira anterior neuroectoderm (ANE) PCS105 cDNA
 library Xenopus laevis cDNA clone XL491g22ex 5', mRNA sequence.
ACCESSION BP707553
VERSION BP707553.1 GI:46055952
KEYWORDS EST.
SOURCE Xenopus laevis (African clawed frog)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 Xenopodinae; Xenopus; Xenopus.
REFERENCE 1 (bases 1 to 631)
 Osada,S., Kitayama,A., Ueno,N. and Taira,M.

TITLE Expression analysis of genes which are expressed in the anterior
 neuroectoderm of Xenopus embryos
JOURNAL Unpublished (2004)
COMMENT Contact: Masanori Taira
 Department of Biological Sciences
 Graduate School of Science, University of Tokyo, Japan
 Science and Technology Corporation, Japan
 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan
 Tel: 81-03-5841-4434
 Fax: 81-03-5841-4434
 Email: m_taira@biol.s.u-tokyo.ac.jp,
 URL: http://www.shigen.nig.ac.jp/ndbp/xenopus/est/.
FEATURES Location/Qualifiers
 1..631
 /organism="Xenopus laevis"
 /mol_type="mRNA"
 /db_xref="taxon:8355"
 /clone="XL490g22ex"
 /issue_type="anterior neuroectoderm"
 /dev_stage="late gastrula (stage 12.5)"
 /clone_lib="Osada Taira anterior neuroectoderm (ANE)
 PCS105 cDNA library"
ORIGIN
 Query Match 70.7%; Score 19.8; DB 3; Length 637;
 Best Local Similarity 91.3%; Pred. No. 8.2e+02;
 Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 3 ACACGATACACAAACCGTTGT 25
 |||||
 |||||
Db 215 ACACGATACACAAACCGTTGT 237
 |||||
 |||||
RESULT 18
LOCUS BP700663 637 bp mRNA linear EST 19-JUL-2004
DEFINITION BP700663 Osada Taira anterior neuroectoderm (ANE) PCS105 cDNA
 library Xenopus laevis cDNA clone XL490g22ex 5', mRNA sequence.
ACCESSION BP700663
VERSION BP700663.1 GI:46049019
KEYWORDS EST.
SOURCE Xenopus laevis (African clawed frog)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 Xenopodinae; Xenopus; Xenopus.
REFERENCE 1 (bases 1 to 637)
 Osada,S., Kitayama,A., Ueno,N. and Taira,M.
 Expression analysis of genes which are expressed in the anterior
 neuroectoderm of Xenopus embryos
 Unpublished (2004)
 Contact: Masanori Taira
 Department of Biological Sciences
 Graduate School of Science, University of Tokyo, Japan
 Science and Technology Corporation, Japan
 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan
 Tel: 81-03-5841-4434
 Fax: 81-03-5841-4434
 Email: m_taira@biol.s.u-tokyo.ac.jp,
 URL: http://www.shigen.nig.ac.jp/ndbp/xenopus/est/.
FEATURES Location/Qualifiers
 1..637
 /organism="Xenopus laevis"
 /mol_type="mRNA"
 /db_xref="taxon:8355"
 /clone="XL490g22ex"
 /issue_type="anterior neuroectoderm"
 /dev_stage="late gastrula (stage 12.5)"
 /clone_lib="Osada Taira anterior neuroectoderm (ANE)
 PCS105 cDNA library"
ORIGIN
 Query Match 70.7%; Score 19.8; DB 3; Length 637;

Best Local Similarity 91.3%; Pred. No. 8.2e+02;
Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 ACAGCAATACAAACCGTTGT 25
Db 215 ACAACAAATACAAACCGTTGT 237

RESULT 19

CO382721/c

LOCUS CO382721 811 bp mRNA linear EST 30-JUN-2004
DEFINITION AGENCOURT 26189827 Blumberg Cho dorsal blastopore 11p Xenopus laevis cDNA clone IMAGE:7295911 5', mRNA sequence.

ACCESSION CO382721
VERSION CO382721.1 GI:49490823

KEYWORDS EST.
SOURCE Xenopus laevis (African clawed frog)

ORGANISM Xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
Xenopodinae; Xenopus; Xenopus.

1 (bases 1 to 811)
NIH-MGC http://mgc.nci.nih.gov/.

JOURNAL Unpublished (1999)
Contact: Daniela S. Gerhard, Ph.D.

Office of Cancer Genomics
National Cancer Institute / NIH

Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgaaps-r@mail.nih.gov

Tissue Procurement: Bruce Blumberg
CDNA Library Preparation: B. Blumberg

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LML)
DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LML at:

http://image.llnl.gov

Plate: LLM15318 row: n column: 05
High quality sequence spot: 639.

Location/Qualifiers

1. 811
/organism="Xenopus laevis"

/mol_type="mRNA"

/db_xref="taxon:8155"

/clone="IMAGE:7295911"

/sex="Both"

/tissue_type="dorsal blastopore 11p"

/lab_host="TOP10"

/clone_lib="Blumberg Cho dorsal blastopore 11p"

/note="Organ: embryo; Vector: pBluescript SK-; Site_1:
ECORI; Site_2: XhoI; Library: prepared from 50 ug of

total RNA by oligo-dT priming and AMV reverse
transcriptase. After addition of EcoRI linkers and

ECORI-XhoI digestion, the cDNA was size selected by
chromatography on Sepharose CL-4B columns and fractions

containing cDNAs larger than 500 bp were ligated into
EcoRI-XhoI-digested lambda ZAPIT (Unizap-XR) and packaged

in vitro. Average insert size is 1.4 kb. The original
library contained 6 x 10⁶ recombinants, of which 3 x 10⁶

were amplified and stored at -70 C in 5M buffer containing
7% DMSO. 3 x 10⁶ pfu were mass excised and the resulting

phagemids used to infect TOP10P. References: Science 253,
196-196 and Methods in Molecular Biology 97, 555-574.

Additional sequences from this library have been deposited
under the name Xenopus laevis dorsal blastopore 11p.

Library constructed by Bruce Blumberg (University of
California, Irvine, Department of Developmental and Cell

Biology)."

ORIGIN

Query Match 70.7%; Score 19.8; DB 8; Length 811;
Best Local Similarity 91.3%; Pred. No. 8.4e+02;
Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 ACAGCAATACAAACCGTTGT 25
Db 146 ACAACAAATACAAACCGTTGT 124

RESULT 20

CN471390

LOCUS CN471390 311 bp mRNA linear EST 22-APR-2004
DEFINITION hh.Ab.Brain2000.000005240 A. burtoni Cichlid Brain cDNA library

2000, Russell Fernald Astacotilapia burtoni cDNA clone
hh.Ab.Brain2000_000005240 5', mRNA sequence.

ACCESSION CN471390
VERSION CN471390.1 GI:46491834

KEYWORDS EST.
SOURCE Astacotilapia burtoni

ORGANISM Astacotilapia burtoni
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percormorpha; Perciformes;
Labroidae; Cichlidae; African cichlids; Pseudocrenilabrinae;

Haplochromini; Astacotilapia.
1 (bases 1 to 311)

Remm, S.C., Aubin-Horth, N., and Hofmann, H.A.
Biologically meaningful expression profiling across species using

heterologous hybridization to a cDNA microarray
BMC Genomics 5 (1), 42 (2004)

JOURNAL 15238158
Contact: Hans A. Hofmann

Bauer Center for Genomics Research
Harvard University
7 Divinity Avenue, Cambridge, MA 02138, USA

Tel: 617-384-8058
Fax: 617-495-2196

Email: hansen@harvard.edu
PCR Primers

FORWARD: C5VP3: AAGCGGCAATTAACCTCACTA
BACKWARD: C5VP2: TTCCAGTCACGACGTTGTAATA

Plate: 054 row: E column: 05
Seq primer: C5VP3: AAGCGGCAATTAACCTCACTA.

Location/Qualifiers

1. 311
/organism="Astacotilapia burtoni"

/mol_type="mRNA"

/db_xref="taxon:8153"

/clone="hh.Ab.Brain2000_000005240"

/sex="mixed"

/tissue_type="brain"

/dev_stage="mixed"

/lab_host="X11-Blue MRF" E.coli"

/clone_lib="A. burtoni Cichlid Brain cDNA library 2000,
Russell Fernald"

/note="Vector: pBS11sk-; Site_1: EcoRI; Site_2: XhoI; mRNA
was isolated from the brains of A. burtoni for both sexes

at all stages of development and reproductive status."

ORIGIN

Query Match 70.0%; Score 19.6; DB 8; Length 311;
Best Local Similarity 84.6%; Pred. No. 9.5e+02;
Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 GAACAGCAATACAAACCGTTGT 26
Db 89 GAACAGCAATACAAACCGTTGT 114

RESULT 21
CB500937/c

LOCUS CB500937 355 bp mRNA linear EST 16-MAY-2003
DEFINITION bsalpbth003037 head Salmo salar cDNA, mRNA sequence.

ACCESSION CB500937
VERSION CB500937.1 GI:29312163

KEYWORDS EST.

ORIGIN

Query Match 70.0%; Score 19.6; DB 8; Length 311;
Best Local Similarity 84.6%; Pred. No. 9.5e+02;
Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

SOURCE Salmo salar (Atlantic salmon)

ORGANISM Salmo salar

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Protacanthopterygii; Salmoniformes; Salmonidae; Salmo. 1 (bases 1 to 355)

AUTHORS Rise, M.L., von Schalburg, K.R., Brown, G.D., Mawer, M.A., Devlin, R.H., Kuipers, N., Busby, M., Beetz-Sargent, M., Alberto, R., Gibbs, A.R., Hunt, P., Shuklin, R., Zelnik, J.A., Nelson, C., Jones, S.R., Smalås, D.E., Jones, S.J., Schein, J.E., Marra, M.A., Butterfield, Y.S., Stott, J.M., Ng, S.H., Davidson, W.S. and Koop, B.F. Development and application of a salmonid EST database and cDNA microarray: data mining and interspecific hybridization characteristics

TITLE Genome Res. 14 (3), 478-490 (2004)

JOURNAL 14962987

COMMENT Contact: Koop, B.F.
Centre for Biomedical Research
University of Victoria
PO Box 3020 STN CSC, Victoria BC, V8W 3N5, Canada
Tel: 250 472 4067
Fax: 250 472 4075
Email: dkoop@uvic.ca
Centre for Biomedical Research, University of Victoria cDNA preparation and sequencing: Roberto Alberto, Marianne Beetz-Sargent, Maura Busby, Peter Hunt, Linda McKinnel, B.F. Koop. bioinformatics: Gordon D Brown.

FEATURES Location/Qualifiers

source 1..355

/organism="Salmo salar"

/mol_type="mRNA"

/strain="McConnell"

/db_xref="taxon:8030"

/clone_lib="head"

/note="Vector: pBluescript1SK+, Library Creator: Kristian R von Schalburg; Atlantic salmon tissue contributors: Carlo Biagi, Mitch Un and Robert Devlin (DFO, Vancouver, B.C.), Simon Jones (PBS, Nanaimo, B.C.), Seaspang Hatchery (Crofton, B.C.), Rachel Roper (University of Victoria)"

ORIGIN

Query Match 70.0%; Score 19.6; DB 4; Length 355;
Best Local Similarity 84.6%; Pred. No. 9.6e+02;
Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 ACACGCAATACACAAACCGTTGTG 28
|||||
283 ACATCATACACAAACCGTTATGTG 258
|||||

RESULT 22 DV085458 487 bp mRNA linear EST 01-DEC-2005

LOCUS 327-384-16 L07 T7 Nematostella vectensis normalized cDNA library

DEFINITION 327 Nematostella vectensis cDNA clone 327-384-16_L07_T7, mRNA

ACCESSION DV085458

VERSION DV085458.1 GI:82866851

KEYWORDS EST.

SOURCE Nematostella vectensis

ORGANISM Eukaryota; Metazoa; Chordata; Anthozoa; Hexacoralia; Actiniaria; Edwardsiidae; Nematostella.

REFERENCE 1 (bases 1 to 487)

AUTHORS Technau, U., Rudd, S., Maxwell, P., Gordon, P.M.K., Saina, M., Grasso, L.C., Hayward, D.C., Sensen, C.W., Saint, R., Holstein, T.W., Ball, E.B. and Miller, D.J.

TITLE Maintenance of ancestral complexity and non-metazoan genes in two basal cnidarians

JOURNAL Trends Genet. (2005) In press

COMMENT Contact: Ulrich Technau
Sars Centre for Marine Molecular Biology

FEATURES Location/Qualifiers

source 1..487

/organism="Nematostella vectensis"

/mol_type="mRNA"

/db_xref="taxon:45351"

/clone="327-384-16_L07_T7"

/sex="thermaphrodite"

/issue_type="whole embryos"

/dev_stage="mixed stages from unfertilized eggs to primary polyps"

/clone_lib="Nematostella vectensis normalized cDNA library 327"

/note="Vector: pBluescript II SK(+); Site_1: NotI; Site_2: EcoRI; Normalized oligo dT primed cDNA library using poly A+ RNA from mixed embryonic stages of Nematostella vectensis. Cloned directionally into pBluescript II SK(+). Ligated in NotI, EcoRI"

ORIGIN

Query Match 70.0%; Score 19.6; DB 10; Length 487;
Best Local Similarity 84.6%; Pred. No. 9.9e+02;
Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 GAACAGCAATACACAAACCGTTGTG 26
|||||
220 GAACCAACATACACAAACCGAAGTG 195
|||||

RESULT 23 BE111725 496 bp mRNA linear EST 13-JUN-2000

LOCUS BE111725

DEFINITION UI-R-BJ1-avx-a-12-0-UI.s1 UI-R-BJ1 Rattus norvegicus cDNA clone

ACCESSION BE111725

VERSION BE111725.1 GI:8503830

KEYWORDS EST.

SOURCE Rattus norvegicus (Norway rat)

ORGANISM Rattus norvegicus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Muridae; Murinae; Rattus.

AUTHORS 1 (bases 1 to 496)

TITLE Bonaldo, M.F., Lemmon, G. and Soares, M.B.

JOURNAL Normalization and subtraction: two approaches to facilitate gene discovery

COMMENT Genome Res. 6 (9), 791-806 (1996)

CONTACT Soares, MB

COORDINATED Laboratory for Computational Genomics

UNIVERSITY of Iowa

375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA

TEL: 319 335 8250

FAX: 319 335 9565

EMAIL: Bento-soares@uiowa.edu

The sequence contained an oligo-dT track that was present in the oligonucleotide that was used to prime the synthesis of first strand cDNA and therefore this may represent a bonafide poly A tail. The sequence tag present in the cDNA between the NotI site and the oligo-dT track served to identify it as a clone from the normalized ventricle at 16.5 dpc library cDNA library Preparation: M.B. Soares Lab Clone distribution: clones will be available through Research Genetics (www.reagen.com)

SEQ PRIMER: M13 Forward

POLYA=Yes.

FEATURES Location/Qualifiers

source 1..496

/organism="Rattus norvegicus"

/mol_type="mRNA"

/strain="Sprague-Dawley"
 /db_xref="taxon:10116"
 /clone="UT-R-BJ1-ava-a-12-0-UT"
 /lab_host="DH10B (Life Technologies)"
 /clone_1ib="01-R-BJ1"
 /note="Vector: pTR73D-Pac1; Site 1: Not I; Site 2: Eco RI;
 The UT-R-BJ1 library is a subtracted library derived from
 the following tissues: atrium at 16.5 dpc, ventricle at
 16.5 dpc, AV canal at 16.5 dpc, atrium at 15 dpc,
 ventricle at 15 dpc, AV canal at 15 dpc, ventricle at 13
 dpc, and adult heart. For a detailed description of the
 library from which this clone was derived, please visit
 our web site at ratec.eng.uci.edu. The subtraction has
 been previously described in (Bonaldo, Lennon and Soares,
 Genome Research 6:791-806, 1996)
 TAG_TISSUE=ventricle at 16.5 dpc
 TAG_LIB=UT-R-BJ1
 TAG_SEQ=GTTCG"

ORIGIN

Query Match 70.0%; Score 19.6; DB 7; Length 496;
 Best Local Similarity 84.6%; Pred. No. 9.9e+02;
 Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 2 AACAGCAATACAAACCGTTGTGT 27
 58 AACGCAATACAAACCGATTGT 83

RESULT 24
 A0325004/c 579 bp DNA linear GSS 08-JAN-1995
 LOCUS mgxb0020B07r CUGI Rice Blast BAC Library Magnaporthe grisea genomic
 DEFINITION clone mgxb0020B07r, genomic survey sequence.
 A0325004
 A0325004.1 GI:4116856

ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Magnaporthe grisea (anamorph: Pyricularia grisea)
 Magnaporthe grisea
 Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 Sordariomycetes incertae sedis; Magnaportheaceae; Magnaporthe.

REFERENCE
 AUTHORS
 Yu, Y., Zhu, H., Boyd, C. A., Gaudette, B., Gayle, A., Kingsbury, R.,
 Phillips, K., Sasinowski, M., Wing, R. A. and Dean, R. A.

TITLE
 A BAC End Sequencing Framework to Sequence the Magnaporthe grisea
 Genome
 JOURNAL
 COMMENT
 Unpublished (1998)
 Contact: Dean RA
 Clemson University Genomics Institute
 100 Jordan Hall, Clemson University, Clemson, SC 29634
 Tel: 864 656 5737
 Fax: 864 656 4293
 Email: rideanc@clemson.edu
 Seq primer: GGAACAAGCTATGACCATG
 Class: BAC ends
 High quality sequence stop: 343.
 Location/Qualifiers

FEATURES
 source
 1..579
 /organism="Magnaporthe grisea"
 /mol_type="genomic DNA"
 /strain="70-15"
 /db_xref="taxon:148305"
 /clone="mgxb0020B07r"
 /issue_type="Protolaests"
 /lab_host="E. coli DH10B"
 /clone_1ib="CUGI Rice Blast BAC Library"
 /note="Vector: pBACWCH; Site 1: HindIII; Site 2: HindIII;
 Rice Blast is one of the most devastating fungal diseases
 of rice world wide. It is a filamentous ascomycete with
 a haploid genome (n=7) of approximately 40 Mbp. Rice
 blast is an important model fungal pathogen for studying
 numerous aspects of the fungal-host interaction. In

order to facilitate genome wide analysis, a BAC library
 containing 9216 clones with an average insert size of 130
 kbp was constructed. This library represents greater
 than 25X genome coverage. High density colony filters
 are available upon request."

ORIGIN

Query Match 70.0%; Score 19.6; DB 11; Length 579;
 Best Local Similarity 84.6%; Pred. No. 1e+03;
 Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 2 AACAGCAATACAAACCGTTGTGT 27
 466 AACGCAATACAAACCGTTGT 441

RESULT 25
 CV232583 658 bp mRNA linear EST 21-SEP-2004
 LOCUS WS0198.B21.A22 PT-DX-N-A-10 Populus trichocarpa cDNA clone
 DEFINITION WS0198.A22.3', mRNA sequence.
 CV232583
 CV232583.1 GI:52388271

ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 EST.
 Populus trichocarpa (Populus balsamifera subsp. trichocarpa)
 Populus trichocarpa
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 rosids; eurosids I; Malpighiales; Salicaceae; Salicaceae; Populus.

REFERENCE
 AUTHORS
 Ralph, S., Cooper, D., Kolosova, N., Oddy, C., Butterfield, Y.,
 Kirkpatrick, R., Liu, J., Palmquist, D., Stott, J., Barber, S., Yang, G.,
 Babalick, R., Brown-John, M., Chand, S., Featherstone, R., Mason, A.,
 Mayo, M., Moran, J., Olson, T., Wong, D., Ritzland, C. E., Siddiqui, A.,
 Holt, R., Jones, S., Marra, M., Ellis, B. E., Douglas, C., Ritzland, K. and
 Bohlmann, J.

TITLE
 The poplar transcriptome: Analysis of expressed sequence tags from
 multiple cDNA libraries
 JOURNAL
 COMMENT
 Unpublished (2004)
 Contact: Joerg Bohlmann
 Genome BC forest genomics program
 University of British Columbia
 Michael Smith Laboratories, 6174 University Boulevard, Rm. 237,
 Vancouver, British Columbia, Canada, V6T 1Z3
 Tel: 1-604-822-0282
 Fax: 1-604-822-2114
 Email: bohlmann@msl.ubc.ca
 Plate: WS0198 row: A column: 22
 High quality sequence stop: 658.
 Location/Qualifiers

FEATURES
 source
 1..658
 /organism="Populus trichocarpa"
 /mol_type="mRNA"
 /cultivar="VF-125"
 /db_xref="taxon:3694"
 /clone="WS0198.A22"
 /sex="Not determined"
 /lab_host="E. coli DH10B T1 phage resistant cells"

/note="Vector: pBluescript II SK (+) XR; Site 1: EcoRI (5'
 end of cDNA); Site 2: XhoI (3' end of cDNA); Outer Xylem
 from 5 year old trees harvested every two weeks between
 April and October of 2002 at the University of British
 Columbia south campus farm in Vancouver, British Columbia.
 mRNA was isolated from each tissue source independently
 and equal quantities of mRNA from each tissue were then
 pooled. cDNA was prepared from 5 micrograms of mRNA and
 directionally ligated into the pBluescript II SK (+) XR
 vector using the pBluescript II XR cDNA library
 construction kit according to manufacturer's instructions
 with modifications (Stratagene). Plasmid DNA was then
 transformed by electroporation into DH10B cells
 (Invitrogen) for propagation. Normalization was applied

ORIGIN according to published methods [Bonaldo M.F. et al. (1996) Genome Research 6(9):791] in order to reduce the abundance of highly expressed transcripts."

Query Match 70.0%; Score 19.6; DB 8; Length 658;
Best Local Similarity 84.6%; Pred. No. 1e+03;
Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 AACGCAATACACAACCGTTGTGT 27
DB 123 AACACCAAAACAAAGAACCTTGTGT 98

RESULT 26
LOCUS CL544775
DEFINITION OB_CL544775.6 OB_Ba Oryza brachyantha genomic clone
OB_BA0073112.5', genomic survey sequence.
CL544775
CL544775.1 GI:47615377
GSS.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

Oryza brachyantha
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Bep
clade; Shrubtoideae; Oryzaceae; Oryza.
1 (bases 1 to 677)
Klm.H., Yu.Y., Stum.D., Yost.D., Rao.K., Luo.M., Jetty.R.,
Kudrna.D., Muller.C., Hatfield.J., Soderlund.C. and Wing.R.
Unpublished (2004)
Contact: Rod A. Wing
Arizona Genomics Institute
University of Arizona
Forbes Building Room 303, Tucson, AZ 85721-0036, USA
Tel: 520 626 9595
Fax: 520 621 1259
Email: rwing@genome.arizona.edu

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

PCR Primers
FORWARD: TAA TAC GAC TCA CTA TAG GG
BACKWARD: CAC TCA TTA GGC ACC CCA
Insert Length: 145000 Std Error: 0.00
Plate: 0073 row: 1 column: 12
Seq primer: TAA TAC GAC TCA CTA TAG GG
Class: BAC ends.
Location/Qualifiers

FEATURES
source

1..677
/organism="Oryza brachyantha"
/mol_type="genomic DNA"
/db_xref="taxon:4533"
/clone="OB_BA0073112"
/issue_type="leaves"
/dev_stage="mature"
/lab_host="DH10B"
/clone_lib="OB_Ba"
/note="Vector: PACIBAC1; Site_1: HindIII; Site_2: HindIII"

ORIGIN
Query Match 70.0%; Score 19.6; DB 13; Length 677;
Best Local Similarity 84.6%; Pred. No. 1e+03;
Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 AACGCAATACACAACCGTTGTGT 27
DB 52 AACGCAATACACAACCGTTGTGT 77

RESULT 27
LOCUS CF234787/c
DEFINITION CF234787 682 bp mRNA linear EST 05-AUG-2003
Populus alba x Populus tremula cDNA 5', mRNA sequence.

ACCESSION CF234787
VERSION CF234787.1 GI:33454216
KEYWORDS
SOURCE
ORGANISM

Populus alba x Populus tremula
Populus alba x Populus tremula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eucosids I; Malpighiales; Salicaceae; Salicaceae; Populus.
1 (bases 1 to 682)
Dejardin,A., Leple,J.-C., Lesage-Descauses,M.-C., Costa,G. and
Pilate,G.

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Expressed sequence tags from poplar wood tissues - A comparative
analysis from multiple libraries
Plant Biol. 6 (1), 55-64 (2004)
15095135
Contact: Leple JC
Unit of Forest Improvement, Genetics and Physiology
National Institute for Agricultural Research (INRA)
Domaine de l'Imere, BP20619 ARDON, 45166 OLIVET CEDEX, FRANCE
Tel: 33 02 38 41 78 00
Fax: 33 02 38 41 78 79
Email: Jean-Charles.Lep@ Orleans.inra.fr
PCR Primers
FORWARD: Triplexa 5' CTCGGGAAGCGCGCATTTGTG 3'
BACKWARD: Triplexa 5' ATACGACTTACTATAGGCGGA 3'
Plate: PtauxT0015 row: B column: 4
Seq primer: Triplexa 5' CTCGGGAAGCGCGCATTTGTG 3'.
Location/Qualifiers

FEATURES
source

1..682
/organism="Populus alba x Populus tremula"
/mol_type="mRNA"
/strain="clone 717-1-B4"
/db_xref="taxon:80863"
/sex="female"
/issue_type="Young differentiating xylem harvested on the
tension wood side of tilted trees"
/dev_stage="3-years-old poplar trees grown in the nursery"
/clone_lib="Poplar cDNA library from young tension xylem"
/note="A cDNA library was made with mRNA isolated from
tension wood tissues corresponding to the differentiating
xylem collected with a scalpel after debarking the stem.
The sampling was done on 3 different tilted trees grown in
the nursery. cDNAs were cloned in an oriented way into
SfiI (A and B) restriction sites. A one-step conversion of
Lambda Triplex2 to the corresponding triplex2 plasmid was
done via site-specific recombination at loxp sites
(Clontech; SMART cDNA library construction kit). cDNA
inserts were PCR amplified using flanking primers and then
sequenced on a ABI3100 Genetic Analyser (Applied
Biosystem)"

ORIGIN

Query Match 70.0%; Score 19.6; DB 5; Length 682;
Best Local Similarity 84.6%; Pred. No. 1e+03;
Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 AACGCAATACACAACCGTTGTGT 27
DB 89 AACGCAATACACAACCGTTGTGT 64

RESULT 28
LOCUS CZ749731
DEFINITION CZ749731 689 bp DNA linear GSS 26-JUL-2005
OC_BA0099B06.r OC_Ba Oryza coarctata genomic clone OC_BA0099B06
3', genomic survey sequence.
CZ749731
CZ749731.1 GI:71185246
GSS.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

Oryza coarctata (Porteresia coarctata)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Bep


```
clade: Ehirhartoidae; Oryzae; Oryza.
REFERENCE
  1 (bases 1 to 689)
  Klm,H., Collura,K., Misotski,M., Byrne,M., Stum,D., Smart,D.,
  Rao,K., Luo,M., Jettly,R., Kudrna,D., Miller,C., Soderlund,C. and
  Wang,R.
  OMAF (Oryza Map Alignment Project) - Arizona Genomics Institute
  Unpublished (2005)
  Contact: Rod A. Wang
  Arizona Genomics Institute
  University of Arizona
  Forbes Building Room 303, Tucson, AZ 85721-0036, USA
  Tel: 520 626 9595
  Fax: 520 621 1259
  Email: rtwing@genome.arizona.edu
PCR Primers
FORWARD: TAA TAC GAC TCA CTA TAG GG
BACKWARD: CAC TCA TTA GGC ACC CCA
Plate: 0099 row: B column: 06
Seq primer: CAC TCA TTA GGC ACC CCA
Class: BAC ends.
FEATURES
  source
    1..689
    /organism="Oryza coarctata"
    /mol_type="genomic DNA"
    /db_xref="taxon:77588"
    /clone="OC_Ba009806"
    /rissue_type="leaves"
    /dev_stage="mature"
    /lab_host="DH10B"
    /clone_1lb="OC_Ba"
    /note="Vector: pAGIBAC1, Site_1: HindIII, Site_2: HindIII"
ORIGIN
  Query Match      70.0%; Score 19.6; DB 13; Length 689;
  Best Local Similarity 84.6%; Pred. No. 1e+03;
  Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY
  2 AACAGCAATACAAACCGTTGTGT 27
  61 AACAGCAATACAAACCGTTGT 86

RESULT 29
  BM085832      691 bp  mRNA  linear  EST 27-MAY-2005
  LOCUS
  DEFINITION
  BM085832 Nori Satoh unpublished cDNA library, larva Clona
  intestinalis cDNA clone rcilv043e07 3', mRNA sequence.
  ACCESSION
  BM085832
  VERSION
  BM085832.1 GI:24261112
  KEYWORDS
  EST.
  SOURCE
  Clona intestinalis
  Clona intestinalis
  Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
  Phlebobranchia; Clonidae; Clona.
  1 (bases 1 to 691)
  Satou,Y., Shin-I,T., Kohara,Y. and Satoh,N.
  Expressed genes in Clona intestinalis (2002c)
  Unpublished (2002)
  Contact: Nori Satoh
  Department of Zoology
  Kyoto University
  Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
  Tel: 81-75-753-4081
  Fax: 81-75-705-1113
  Email: satoh@ascidian.zool.kyoto-u.ac.jp.
  Location/Qualifiers
    1..691
    /organism="Clona intestinalis"
    /mol_type="mRNA"
    /db_xref="taxon:7719"
    /clone="rcilv043e07"
    /rissue_type="whole animal"
    /dev_stage="larva"

FEATURES
  source
    1..689
    /organism="Oryza coarctata"
    /mol_type="genomic DNA"
    /db_xref="taxon:77588"
    /clone="OC_Ba009806"
    /rissue_type="leaves"
    /dev_stage="mature"
    /lab_host="DH10B"
    /clone_1lb="OC_Ba"
    /note="Vector: pAGIBAC1, Site_1: HindIII, Site_2: HindIII"
ORIGIN
  Query Match      70.0%; Score 19.6; DB 13; Length 689;
  Best Local Similarity 84.6%; Pred. No. 1e+03;
  Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY
  2 AACAGCAATACAAACCGTTGTGT 27
  61 AACAGCAATACAAACCGTTGT 86

RESULT 29
  BM085832      691 bp  mRNA  linear  EST 27-MAY-2005
  LOCUS
  DEFINITION
  BM085832 Nori Satoh unpublished cDNA library, larva Clona
  intestinalis cDNA clone rcilv043e07 3', mRNA sequence.
  ACCESSION
  BM085832
  VERSION
  BM085832.1 GI:24261112
  KEYWORDS
  EST.
  SOURCE
  Clona intestinalis
  Clona intestinalis
  Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
  Phlebobranchia; Clonidae; Clona.
  1 (bases 1 to 691)
  Satou,Y., Shin-I,T., Kohara,Y. and Satoh,N.
  Expressed genes in Clona intestinalis (2002c)
  Unpublished (2002)
  Contact: Nori Satoh
  Department of Zoology
  Kyoto University
  Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
  Tel: 81-75-753-4081
  Fax: 81-75-705-1113
  Email: satoh@ascidian.zool.kyoto-u.ac.jp.
  Location/Qualifiers
    1..691
    /organism="Clona intestinalis"
    /mol_type="mRNA"
    /db_xref="taxon:7719"
    /clone="rcilv058d20"
    /rissue_type="whole animal"
    /dev_stage="larva"
    /clone_1lb="Nori Satoh unpublished cDNA library, larva"

ORIGIN
  Query Match      70.0%; Score 19.6; DB 3; Length 735;
  Best Local Similarity 84.6%; Pred. No. 1e+03;
  Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY
  2 AACAGCAATACAAACCGTTGTGT 27
  87 AACAGCAATACAAACCGTTGT 112

RESULT 31
  DM577165      737 bp  mRNA  linear  EST 17-JAN-2006
  LOCUS
  DEFINITION
  DM577165 Nori Satoh unpublished cDNA library, larva Clona
  intestinalis cDNA clone rcilv058d20 3', mRNA sequence.
  ACCESSION
  DM577165
  VERSION
  DM577165.1 GI:85048987
  KEYWORDS
  EST.
  SOURCE
  Salmo salar (Atlantic salmon)
  Salmo salar
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Actinopterygii; Neopterygii; Teleostei; Euteleostei;
  Protacanthopterygii; Salmoniformes; Salmonidae; Salmo.
  1 (bases 1 to 737)
  Koop,B.P., Davidson,W.S. and CGRAP Consortium.
  Expressed Sequence Tags from Atlantic Salmon Unpublished (2006)
  Unpublished (2006)
  Contact: Koop BP

REFERENCE
  AUTHORS
  TITLE
  JOURNAL
  COMMENT
```

Centre for Biomedical Research
University of Victoria
PO Box 3020 STN CSC, Victoria BC, V8W 3N5, Canada
Tel: 250 472 4067
Fax: 250 472 4075
Email: bkoop@vic.ca
Sequenced by: Genome Sciences Centre, BC Cancer Agency Rob Holt, M. Marra. Bioinformatics: Centre for Biomedical Research, University of Victoria Jong Leong, BF Koop.
Insert Length: 737 Std Error: 0.00
Plate: 567
Seq primer: M13 Forward
High quality sequence stop: 737.
Location/Qualifiers
1. 737
/organism="Salmo salar"
/mol_type="mRNA"
/strain="McConnell"
/db_xref="taxon:8030"
/clone="asal_rgd_567_161_fwd"
/issue_type="mixed tissue"
/clone_lib="rgd2"
/note="Organ: brain, kidney, spleen; Vector: pCMVSPORT6; asalrgb2 mixed tissue Salmo salar cDNA, mRNA sequence; contributors: Robert Devlin (DFO, Vancouver, B.C.)"

ORIGIN
Query Match 70.0%; Score 19.6; DB 10; Length 737;
Best Local Similarity 84.6%; Pred. No. 1e+03; Mismatches 22; Conservative 0; Indels 0; Gaps 0;

QY 3 ACAGCAATACACAAACCGTTGTGTG 28
Db 692 ACATCAATACACACCCCTTATGTG 717

RESULT 32
CA064299 753 bp mRNA linear EST 04-MAR-2003
LOCUS asalrgb530194 mixed tissue Salmo salar cDNA, mRNA sequence.
DEFINITION CA064299
ACCESSION CA064299
VERSION CA064299.1 GI:24394542
KEYWORDS EST.
SOURCE Salmo salar (Atlantic salmon)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Protacanthopterygii; Salmoniformes; Salmonidae; Salmo.
1 (bases 1 to 753)
Rise,M.L., von Schalburg,K.R., Brown,G.D., Mawer,M.A., Devlin,R.H., Kuipers,N., Busby,M., Beetz-Sargent,M., Alberto,R., Gibbs,A.R., Hunt,P., Shukin,R., Zelnik,J.A., Nelson,C., Jones,S.R., Smalhus,D.E., Jones,S.J., Schein,J.E., Marra,M.A., Butterfield,Y.S., Stott,J.M., Ng,S.H., Davidson,W.S. and Koop,B.F. Development and application of a salmonid EST database and cDNA microarray: data mining and interspecific hybridization characteristics
Genome Res. 14 (3), 478-490 (2004)
14962987
JOURNAL PUBMED
COMMENT Contact: Koop BF
Centre for Biomedical Research
University of Victoria
PO Box 3020 STN CSC, Victoria BC, V8W 3N5, Canada
Tel: 250 472 4067
Fax: 250 472 4075
Email: bkoop@vic.ca
Genome Sciences Centre, BC Cancer Agency
CDNA preparation, sequencing and bioinformatics:
Y Butterfield, R Kirkpatrick, J Asano, N Gilm, R Guin, D Lee, S Lee, T Olson, P Pandoh, A Prabhu, D Smalhus, L Spence, J Stott, S Taylor, G Yang, J Schein, S Jones and M Marra.
POLYA=Yes.
Location/Qualifiers

FEATURES
source

source
1. 753
/organism="Salmo salar"
/mol_type="mRNA"
/strain="McConnell"
/db_xref="taxon:8030"
/clone_lib="mixed tissue"
/note="Vector: pCMVSPORT6; Library Creator: Research Genetics; Atlantic salmon tissue contributors: Carlo Biagi, Mitch Un and Robert Devlin (DFO, Vancouver, B.C.), Simon Jones (PBS, Nanaimo, B.C.), Seaspining Hatchery (Crofton, B.C.), Rachel Roper (University of Victoria)"

ORIGIN
Query Match 70.0%; Score 19.6; DB 4; Length 753;
Best Local Similarity 84.6%; Pred. No. 1e+03; Mismatches 22; Conservative 0; Indels 0; Gaps 0;

QY 3 ACAGCAATACACAAACCGTTGTGTG 28
Db 670 ACATCAATACACACCCCTTATGTG 695

RESULT 33
CA056575 764 bp mRNA linear EST 04-MAR-2003
LOCUS asalrgb536212 mixed tissue Salmo salar cDNA, mRNA sequence.
DEFINITION CA056575
ACCESSION CA056575
VERSION CA056575.1 GI:24386818
KEYWORDS EST.
SOURCE Salmo salar (Atlantic salmon)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Protacanthopterygii; Salmoniformes; Salmonidae; Salmo.
1 (bases 1 to 764)
Rise,M.L., von Schalburg,K.R., Brown,G.D., Mawer,M.A., Devlin,R.H., Kuipers,N., Busby,M., Beetz-Sargent,M., Alberto,R., Gibbs,A.R., Hunt,P., Shukin,R., Zelnik,J.A., Nelson,C., Jones,S.R., Smalhus,D.E., Jones,S.J., Schein,J.E., Marra,M.A., Butterfield,Y.S., Stott,J.M., Ng,S.H., Davidson,W.S. and Koop,B.F. Development and application of a salmonid EST database and cDNA microarray: data mining and interspecific hybridization characteristics
Genome Res. 14 (3), 478-490 (2004)
14962987
JOURNAL PUBMED
COMMENT Contact: Koop BF
Centre for Biomedical Research
University of Victoria
PO Box 3020 STN CSC, Victoria BC, V8W 3N5, Canada
Tel: 250 472 4067
Fax: 250 472 4075
Email: bkoop@vic.ca
Genome Sciences Centre, BC Cancer Agency
CDNA preparation, sequencing and bioinformatics:
Y Butterfield, R Kirkpatrick, J Asano, N Gilm, R Guin, D Lee, S Lee, T Olson, P Pandoh, A Prabhu, D Smalhus, L Spence, J Stott, S Taylor, G Yang, J Schein, S Jones and M Marra.
Location/Qualifiers
1. 764
/organism="Salmo salar"
/mol_type="mRNA"
/strain="McConnell"
/db_xref="taxon:8030"
/clone_lib="mixed tissue"
/note="Vector: pCMVSPORT6; Library Creator: Research Genetics; Atlantic salmon tissue contributors: Carlo Biagi, Mitch Un and Robert Devlin (DFO, Vancouver, B.C.), Simon Jones (PBS, Nanaimo, B.C.), Seaspining Hatchery (Crofton, B.C.), Rachel Roper (University of Victoria)"

ORIGIN
Query Match 70.0%; Score 19.6; DB 4; Length 764;
Best Local Similarity 84.6%; Pred. No. 1e+03; Mismatches 22; Conservative 0; Indels 0; Gaps 0;

Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 3 ACAGCATACAAACCGTTGTG 28
 |||||
 DB 686 ACATCATACAAACCCCTTATGT 711

RESULT 34
 CB643162/c
 LOCUS
 DEFINITION OSJNB03108, r OSJNB Oryza sativa (japonica cultivar-group) cDNA
 ACCESSION CB643162
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

782 bp mRNA linear EST 08-APR-2003
 Oryza sativa (japonica cultivar-group)
 Oryza sativa (japonica cultivar-group)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; BEP
 clade; Ehrhartoideae; Oryzaceae; Oryza.
 1 (bases 1 to 782)
 Jantsuriyarat C., Gowda, M., Haller, K., Hatfield, J., Lu, G.,
 Stahlberg, E., Zhou, B., Li, H., Kim, H., Yu, Y., Dean, R. A., Wing, R. A.,
 Soderlund, C. and Wang, G. L.
 Large-scale identification of expressed sequence tags involved in
 rice and rice blast fungus interaction
 Plant Physiol. 138 (1), 105-115 (2005)
 1588683
 Contact: Rod Wing
 Arizona Genomics Institute
 University of Arizona
 Biological Sciences West, 440A, P.O. Box 210088, Tucson, AZ
 85721-0088, USA
 Tel: 520 626 3967
 Fax: 520 621 9288
 Email: twing@genome.arizona.edu
 PCR Primers
 FORWARD: gta aac cga cgg cca gtc
 BACKWARD: gga aac agc tac gac cat g
 Plate: 03 row: L column: 08
 Seq primer: gga aac agc tac gac cat g.
 Location/Qualifiers
 1. 782
 /organism="Oryza sativa (japonica cultivar-group)"
 /mol_type="mRNA"
 /cultiyar="Nipponbare"
 /db_xref="taxon:39947"
 /clone="OSJNB03108"
 /issue_type="Leaf"
 /dev_stage="3 week"
 /lab_host="DH10B"
 /clone_lib="OSJNB"
 /note="Vector: Bluescript II KS+; Site 1: EcoRI; Site 2:
 XhoI; 24 hrs after inoculation with Rice Blast (Che
 86061)"

ORIGIN

Query Match 70.0%; Score 19.6; DB 4; Length 782;
 Best Local Similarity 84.6%; Pred. No. 1e+03;
 Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 2 AACGCAATACAAACCGTTGTG 27
 |||||
 DB 29 AAAAGCAATGCAACCAACCGTTGTG 4

RESULT 35
 BU255194
 LOCUS
 DEFINITION BU255194, 792 bp mRNA linear EST 26-NOV-2002
 60374612cF1 CSEQCHN38 Gallus gallus cDNA clone CHEST651c8 5', mRNA
 sequence.
 ACCESSION BU255194
 VERSION BU255194.1 GI:25511921

KEYWORDS EST.
 SOURCE Gallus gallus (chicken)
 ORGANISM Gallus gallus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
 Phasianinae; Gallus.

REFERENCE
 AUTHORS Boardman, P. E., Sanz-Ezquerro, J., Overton, I. M., Burt, D. W., Bosch, E.,
 Pong, W. T., Tickle, C., Brown, W. R. A., Wilson, S. A. and Hubbard, S. J.
 A Comprehensive Collection of Chicken CDNA
 Curr. Biol. 12 (22), 1965-1969 (2002)
 12445392
 Contact: Simon Hubbard
 Department of Biomolecular Sciences
 University of Manchester Institute of Science and Technology
 (UMIST)
 PO Box 88, Manchester, M60 10D, UK
 Tel: 01612008930
 Fax: 01612360409
 Email: Simon.Hubbard@umist.ac.uk.

FEATURES
 source
 1. 792
 /organism="Gallus gallus"
 /mol_type="mRNA"
 /strain="White Leghorn, H1sex"
 /db_xref="taxon:9031"
 /clone="CHEST651c8"
 /dev_stage="36"
 /lab_host="DH10B"
 /clone_lib="CSEQCHN38"
 /note="Organ: limbs; Vector: pbluescript II KS(+); Site 1:
 EcoRI; Site 2: NotI; This normalized library was
 constructed from 1 million independent clones. cDNA
 synthesis was initiated using an oligo(dT) primer, using
 methylated C in the first strand synthesis reaction.
 Following this first strand reaction, double-stranded cDNA
 was bluntend, ligated to NotI adapters, digested with
 EcoRI, size-selected, and cloned into the NotI and EcoRI
 compatible sites of a custom modified MCS of the
 pbluescript (KS+) vector. The library was normalized in 2
 rounds using conditions adapted from Soares et al., PNAS
 (1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6
 (1996): 791, except that a significantly longer
 reannealing hybridization was used."

ORIGIN

Query Match 70.0%; Score 19.6; DB 3; Length 792;
 Best Local Similarity 84.6%; Pred. No. 1e+03;
 Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 3 ACAGCATACAAACCGTTGTG 28
 |||||
 DB 652 ACAGCATACAAACCGTTGTG 627

RESULT 36
 BU087205
 LOCUS BU087205
 DEFINITION BU087205 Nori Satoh unpublished cDNA library, larva Clona
 intestinalis cDNA clone rcily048h03 3', mRNA sequence.
 ACCESSION BU087205
 VERSION BU087205.1 GI:24262485
 KEYWORDS
 SOURCE
 ORGANISM

Clona intestinalis
 Clona intestinalis
 Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
 Phlebobranchia; Cloniidae; Clona.
 1 (bases 1 to 793)
 Satou, Y., Shin-I, T., Kohara, Y. and Satoh, N.
 Expressed genes in Clona intestinalis (2002c)
 Unpublished (2002)
 Contact: Nori Satoh
 Department of Zoology

Kyoto University
Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
Tel: 81-75-753-4081
Fax: 81-75-705-1113
Email: sato@acidian.zool.kyoto-u.ac.jp.

FEATURES

source

Location/Qualifiers
1. .793
/organism="Clona intestinalis"
/mol_type="mRNA"
/db_xref="taxon:7719"
/clone="rc1y048h03"
/tissue_type="whole animal"
/dev_stage="larva"
/clone_lib="Norl Satoh unpublished cDNA library, larva"

ORIGIN

Query Match 70.0%; Score 19.6; DB 3; Length 793;
Best Local Similarity 84.6%; Pred. No. 1e+03; Indels 0; Gaps 0;
Matches 22; Conservative 0; Mismatches 4;

Qy 2 AACGACATACAAACCGTTGTG 27
Db 99 AACGACATACAAACCGTTGTT 124

RESULT 37
DM577166 822 bp mRNA linear EST 17-JAN-2006
LOCUS
DEFINITION
EST_ssal_rpb2_41585_rpb2_Salmo salar clone
DM577166
VERSION
KEYWORDS
SOURCE
ORGANISM
Salmo salar (Atlantic salmon)
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei;
Protacanthopterygii; Salmoniformes; Salmonidae; Salmo.
1 (bases 1 to 822)
Koop, B.F., Davidson, W.S. and CGRAP Consortium.
Expressed Sequence Tags from Atlantic Salmon Unpublished (2006)
Unpublished (2006)
Contact: Koop BF
Centre for Biomedical Research
University of Victoria
PO Box 3020 STN CSC, Victoria BC, V8W 3N5, Canada
Tel: 250 472 4067
Fax: 250 472 4075
Email: bkoop@uvic.ca

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Unpublished Sequence Tags from Atlantic Salmon Unpublished (2006)
Expressed Sequence Tags from Atlantic Salmon Unpublished (2006)
Unpublished (2006)
Contact: Koop BF
Centre for Biomedical Research
University of Victoria
PO Box 3020 STN CSC, Victoria BC, V8W 3N5, Canada
Tel: 250 472 4067
Fax: 250 472 4075
Email: bkoop@uvic.ca

Sequenced by: Genome Sciences Centre, BC Cancer Agency Rob Holt, M.
Marrs, Bioinformatics; Centre for Biomedical Research, University
of Victoria, Jong Leong, BF Koop.
Insert Length: 822 Std Error: 0.00
Plate: 567
Seq primer: M13 Reverse
High quality sequence stop: 822.

FEATURES

source

Location/Qualifiers
1. .822
/organism="Salmo salar"
/mol_type="mRNA"
/strain="McConnell"
/db_xref="taxon:8030"
/clone="asal_rpb2_567_161_rev"
/tissue_type="mixed tissue"
/clone_lib="rpb2"
/note="Organ: brain, kidney, spleen; Vector: pCMVSPORT6;
ssalrpb2 mixed tissue Salmo salar cDNA; Tissue
contributors: Robert Devlin (DFO, Vancouver, B.C.)"

ORIGIN

Query Match 70.0%; Score 19.6; DB 10; Length 822;
Best Local Similarity 84.6%; Pred. No. 1e+03; Indels 0; Gaps 0;
Matches 22; Conservative 0; Mismatches 4;

Qy 3 ACAGCATACAAACCGTTGTG 28
Db 712 ACATCATACAAACCGCTTATGTG 687

RESULT 38
CG219751 832 bp DNA linear GSS 22-AUG-2003
LOCUS
DEFINITION
CGWLQ26TH_ZM_0.7_1.5_KB Zea mays genomic clone ZMMBMA0612P03,
genomic survey sequence.
CG219751
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Zea mays
Zea mays
Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 832)
Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,
Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.
Consortium for Maize Genomics
Unpublished (2002)
Other GSSs: CGWLQ26TV
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TR
Class: methylation filtered.
Location/Qualifiers
1. .832
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBMA0612P03"
/clone_lib="ZM_0.7_1.5_KB"
/note="Vector: pBCKS-; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Unpublished (2002)
Other GSSs: CGWLQ26TV
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TR
Class: methylation filtered.
Location/Qualifiers
1. .832
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBMA0612P03"
/clone_lib="ZM_0.7_1.5_KB"
/note="Vector: pBCKS-; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

FEATURES
source
Location/Qualifiers
1. .832
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBMA0612P03"
/clone_lib="ZM_0.7_1.5_KB"
/note="Vector: pBCKS-; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

ORIGIN

Query Match 70.0%; Score 19.6; DB 12; Length 832;
Best Local Similarity 84.6%; Pred. No. 1e+03; Indels 0; Gaps 0;
Matches 22; Conservative 0; Mismatches 4;

Qy 1 GAACGACATACAAACCGTTGTG 26
Db 676 GAACGACATACATTAACAGTTTGG 651

RESULT 39
CG452337 839 bp DNA linear GSS 17-SEP-2003
LOCUS
DEFINITION
CGVHQ26TV_ZM_0.7_1.5_KB Zea mays genomic clone ZMMBMA0485F03,
genomic survey sequence.
CG452337
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Zea mays
Zea mays
Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 839)
Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,
Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,
Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.
Consortium for Maize Genomics

JOURNAL Unpublished (2002)
 COMMENT Other GSS: OGVH026TH
 CONTACT: Cathy Whitelaw
 TIGR
 9712 Medical Center Drive, Rockville, MD 20850, USA
 Tel: 301-838-5843
 Fax: 301-838-0208
 Email: whitelaw@tigr.org
 Seq primer: TP
 Class: methylation filtered.
 Location/Qualifiers
 1. 839
 /organism="Zea mays"
 /mol_type="genomic DNA"
 /strain="B73"
 /db_xref="taxon:4577"
 /clone_1db="ZMMBMA0485F03"
 /clone_1lb="ZM 0.7 1.5 KB"
 /note="Vector: pBCK-; Site 1: HindIII; 0.7-1.5 kb
 methylation filtered genomic DNA library"

ORIGIN
 Query Match 70.0%; Score 19.6; DB 12; Length 839;
 Best Local Similarity 84.6%; Pred. No. 1e+03;
 Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 GAACAGCAATACAAACCGTTGTG 26
 |||||
 DB 720 GAACAGCAAAACATTAACAGTTTGG 745

RESULT 40
 CG219765 959 bp DNA linear GSS 22-AUG-2003
 LOCUS OGVH026TV ZM 0.7 1.5 KB Zea mays genomic clone ZMMBMA0612F03,
 DEFINITION genomic survey sequence.
 ACCESSION CG219765
 VERSION CG219765.1 GI:34119653
 KEYWORDS GSS.
 SOURCE Zea mays
 ORGANISM Zea mays
 Zea mays
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoideae; Andropogoneae; Zea.
 1 (bases 1 to 959)
 WhiteIaw,C.A., Quackenbush,J., Van Aken,S., Uterback,T.,
 Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
 Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
 Consortium for Maize Genomics
 Unpublished (2002)
 Other GSSs: OGVH026TH
 CONTACT: Cathy Whitelaw
 TIGR
 9712 Medical Center Drive, Rockville, MD 20850, USA
 Tel: 301-838-5843
 Fax: 301-838-0208
 Email: whitelaw@tigr.org
 Seq primer: TP
 Class: methylation filtered.
 Location/Qualifiers
 1. 959
 /organism="Zea mays"
 /mol_type="genomic DNA"
 /strain="B73"
 /db_xref="taxon:4577"
 /clone_1db="ZMMBMA0612F03"
 /clone_1lb="ZM 0.7 1.5 KB"
 /note="Vector: pBCK-; Site 1: HindIII; 0.7-1.5 kb
 methylation filtered genomic DNA library"

ORIGIN
 Query Match 70.0%; Score 19.6; DB 12; Length 959;
 Best Local Similarity 84.6%; Pred. No. 1e+03;
 Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 GAACAGCAATACAAACCGTTGTG 26
 |||||
 DB 737 GAACAGCAAAACATTAACAGTTTGG 762

RESULT 41
 CZ104695 595 bp DNA linear GSS 13-JAN-2005
 LOCUS OM_Ba0113L07.r OM_Ba Oryza minuta genomic clone OM_Ba0113L07.3,
 DEFINITION genomic survey sequence.
 ACCESSION CZ104695
 VERSION CZ104695.1 GI:57680996
 KEYWORDS GSS.
 SOURCE Oryza minuta
 ORGANISM Oryza minuta
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; BAP
 clade; Ehrhartoideae; Oryzaceae; Oryza.
 1 (bases 1 to 595)
 Kim,H., Collura,K., Wisotscki,M., Byrne,M., Stum,D., Smart,D.,
 Rao,K., Luo,M., Jetty,R., Kudrna,D., Muller,C., Soderlund,C. and
 Wang,R.
 ONAP (Oryza Map Alignment Project) - Arizona Genomics Institute
 Unpublished (2005)
 CONTACT: Rod A. Wang
 Arizona Genomics Institute
 University of Arizona
 Forbes Building Room 303, Tucson, AZ 85721-0036, USA
 Tel: 520 626 9595
 Fax: 520 621 1259
 Email: rtwing@genome.arizona.edu
 Plate: 0113 row: L column: 07
 Class: BAC ends.
 Location/Qualifiers
 1. 595
 /organism="Oryza minuta"
 /mol_type="genomic DNA"
 /db_xref="taxon:63629"
 /clone_1db="OM_Ba0113L07"
 /issue_type="leaves"
 /lab_host="DH10B"
 /clone_1lb="OM_Ba"
 /note="Vector: pCUGIBAC1; Site_1: HindIII; Site_2:
 HindIII"

ORIGIN
 Query Match 69.3%; Score 19.4; DB 13; Length 595;
 Best Local Similarity 95.2%; Pred. No. 1.2e+03;
 Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 7 CAATACAAACAAACCGTTGTG 27
 |||||
 DB 203 CAATACAAACAAACCGTTGTG 223

RESULT 42
 DE039251 522 bp DNA linear GSS 16-AUG-2005
 LOCUS DE039251 Branchiostoma floridae DNA, clone: CH302-061024.F, genomic survey
 DEFINITION sequence.
 ACCESSION DE039251
 VERSION DE039251.1 GI:62283060
 KEYWORDS GSS.
 SOURCE Branchiostoma floridae (Florida lancelet)
 ORGANISM Branchiostoma floridae
 Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae;
 Branchiostoma.
 1
 Fujiyama,A., Toyoda,A., Kuroki,Y. and Sakaki,Y.
 BAC end sequences of CHORI-302 Amphioxus library
 Published Only in Database (2005)

```

REFERENCE      2 (bases 1 to 522)
AUTHORS        Fujiyama, A.
TITLE          Direct Submision
JOURNAL        Submitted (31-MAR-2005) Asao Fujiyama, The Institute of Physical
              and Chemical Research (RIKEN), Genomic Sciences Center (GSC),
              1-7-22 Suenhiro-Chou, Tsurumi-Ku, Yokohama, Kanagawa, 230-0045, Japan
              (E-mail:afujiyam@gsc.riken.jp, URL:http://hsp.gsc.riken.jp/,
              Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT        PRIMERS
              Sequencing : T7
LIBRARY        Vector : pTARBAC2.1
              R.site 1 : EcoRI
              R.site 2 : EcoRI.
FEATURES       Location/Qualifiers
              1..522
                /organism="Branchiostoma floridae"
                /mol_type="genomic DNA"
                /db_xref="taxon:7739"
                /clone="CH302-061024.F"
                /sex="male"
                /issue_type="sperm"
                /clone_lib="CHOR10302 Amphioxus genomic BAC library"
                /note="common name:amphioxus"

ORIGIN
Query Match      68.6%; Score 19.2; DB 14; Length 522;
Best Local Similarity 87.5%; Pred. No. 1.5e+03;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY              3 ACAGCAATACACAAACCGTTGTG 26
Db              270 ACAGCAATACATCATTCATTGTG 293

RESULT 43
B0626720      524 bp mRNA linear EST 01-JUL-2002
LOCUS         p226b03.y1 Pratylenchus penetrans mixed stage SL1 TOPO v1
DEFINITION    Pratylenchus penetrans cDNA 5' similar to SW:RS15_HUMAN P11174 40S
RIBOSOMAL     RIBOSOMAL PROTEIN S15 ;, mRNA sequence.
VERSION       B0626720 GI:21653898
KEYWORDS      EST.
SOURCE        Pratylenchus penetrans
ORGANISM      Pratylenchus penetrans
              Eukaryota; Metazoa; Nematoda; Chromadorea; Tylenchida; Tylenchina;
              Tylenchoidea; Pratylenchidae; Pratylenchinae; Pratylenchus.
REFERENCE     1 (bases 1 to 524)
AUTHORS       McCarter, J., Clifton, S., Chiapelli, B., Pape, D., Martin, J.,
              Wylie, T., Dante, M., Marra, M., Hillier, L., Kucaba, T., Theising, B.,
              Bowers, Y., Gibbons, M., Ritter, E., Bennett, J., Franklin, C.,
              Tsagarishevili, R., Ronko, I., Kennedy, S., Maguire, L., Beck, C.,
              Underwood, K., Steptoe, M., Allen, M., Person, B., Swaller, T.,
              Harvey, N., Schurk, R., Kohn, S., Shin, T., Jackson, Y., Cardenas, M.,
              McCann, R., Waterston, R. and Wilson, R.
              The Washington Univ. Nematode EST Project, 1999
              Unpublished (1999)
              Contact: McCarter JP
              The Washington Univ. Nematode EST Project, 1999
              Washington University School of Medicine
              4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
              Tel: 314 286 1800
              Fax: 314 286 1810
              Email: est@wustl.edu
              The library was constructed by Claire Murphy and Dr. James McCarter
              at Washington University, St. Louis, RNA was provided by Andrew
              Kloek of Divergence Inc., St. Louis, MO.
              Putative full length read
              The vector to vector length is 567
              Seq primer: SL1 primer.
FEATURES       Location/Qualifiers
              1..524

```

```

              /organism="Pratylenchus penetrans"
              /mol_type="mRNA"
              /db_xref="taxon:45929"
              /dev_stage="mixed stage"
              /lab_host="DH10B (Invitrogen)"
              /clone_lib="Pratylenchus penetrans mixed stage SL1 TOPO
              v1"
              /note="Vector: PCR11-TOPO (Invitrogen); Site 1: EcoRI;
              Site 2: EcoRI; The library was constructed by Claire
              Murphy and Dr. James McCarter at Washington University,
              St. Louis. Oligo (dT)-SL1 PCR based library. cDNA PCR
              products of size >400 nucleotides containing SL1 on the 5'
              end and oligo (dT) on the 3' end were non-directionally
              cloned into PCR11-TOPO (Invitrogen) following the TOPO TA
              cloning protocol. RNA was provided by Andrew Kloek of
              Divergence, Inc., St. Louis, MO."

ORIGIN
Query Match      68.6%; Score 19.2; DB 3; Length 524;
Best Local Similarity 87.5%; Pred. No. 1.5e+03;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY              1 GAACAGCAATACACAAACCGTTG 24
Db              114 GAGCAGCAATTCACAAACCGTTG 137

RESULT 44
B0627018      524 bp mRNA linear EST 01-JUL-2002
LOCUS         p230b06.y1 Pratylenchus penetrans mixed stage SL1 TOPO v1
DEFINITION    Pratylenchus penetrans cDNA 5' similar to SW:RS15_HUMAN P11174 40S
RIBOSOMAL     RIBOSOMAL PROTEIN S15 ;, mRNA sequence.
VERSION       B0627018 GI:21654196
KEYWORDS      EST.
SOURCE        Pratylenchus penetrans
ORGANISM      Pratylenchus penetrans
              Eukaryota; Metazoa; Nematoda; Chromadorea; Tylenchida; Tylenchina;
              Tylenchoidea; Pratylenchidae; Pratylenchinae; Pratylenchus.
REFERENCE     1 (bases 1 to 524)
AUTHORS       McCarter, J., Clifton, S., Chiapelli, B., Pape, D., Martin, J.,
              Wylie, T., Dante, M., Marra, M., Hillier, L., Kucaba, T., Theising, B.,
              Bowers, Y., Gibbons, M., Ritter, E., Bennett, J., Franklin, C.,
              Tsagarishevili, R., Ronko, I., Kennedy, S., Maguire, L., Beck, C.,
              Underwood, K., Steptoe, M., Allen, M., Person, B., Swaller, T.,
              Harvey, N., Schurk, R., Kohn, S., Shin, T., Jackson, Y., Cardenas, M.,
              McCann, R., Waterston, R. and Wilson, R.
              The Washington Univ. Nematode EST Project, 1999
              Unpublished (1999)
              Contact: McCarter JP
              The Washington Univ. Nematode EST Project, 1999
              Washington University School of Medicine
              4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
              Tel: 314 286 1800
              Fax: 314 286 1810
              Email: est@wustl.edu
              The library was constructed by Claire Murphy and Dr. James McCarter
              at Washington University, St. Louis, RNA was provided by Andrew
              Kloek of Divergence Inc., St. Louis, MO.
              Putative full length read
              The vector to vector length is 576
              Seq primer: SL1 primer.
FEATURES       Location/Qualifiers
              1..524
                /organism="Pratylenchus penetrans"
                /mol_type="mRNA"
                /db_xref="taxon:45929"
                /dev_stage="mixed stage"
                /lab_host="DH10B (Invitrogen)"
                /clone_lib="Pratylenchus penetrans mixed stage SL1 TOPO
                v1"
                /note="Vector: PCR11-TOPO (Invitrogen); Site 1: EcoRI;

```

Site 2: Ecori; The library was constructed by Claire Murphy and Dr. James McCarter at Washington University, St. Louis. Oligo(dT)-SLI PCR based library. cDNA PCR products of size >400 nucleotides containing SLI on the 5' end and oligo(dT) on the 3' end were non-directionally cloned into pCRII-TOPO(Invitrogen) following the Topo TA cloning protocol. RNA was provided by Andrew Kloek of Divergence, Inc., St. Louis, MO."

ORIGIN

Query Match 68.6%; Score 19.2; DB 3; Length 524;
Best Local Similarity 87.5%; Pred. No. 1.5e+03;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GAACAGCAATTCACAAACGCTTG 24
Db 123 GAGCAGCAATTCACAAACGCTTG 146

RESULT 45

BO627107 524 bp mRNA linear EST 01-JUL-2002
LOCUS p31d03.y1 Pratylenchus penetrans mixed stage SLI TOPO v1

DEFINITION Pratylenchus penetrans cDNA 5' similar to SW:RS15_HUMAN P11174 40S
RIBOSOMAL PROTEIN S15 ; , mRNA sequence.

BO627107
BO627107.1 GI:21654285

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

McCarter, J., Clifton, S., Chiapelli, B., Pape, D., Martin, J., Wylie, T., Dante, M., Marra, M., Hillier, L., Kucaba, T., Theising, B., Bowers, Y., Gibbons, M., Ritzer, E., Bennett, J., Franklin, C., Tsagarishevili, R., Ronko, I., Kennedy, S., Maguire, L., Beck, C., Underwood, K., Steptoe, M., Allen, M., Person, B., Swaller, T., Harvey, N., Schurk, R., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.
The Washington Univ. Nematode EST Project, 1999
Unpublished (1999)
Contact: McCarter JP
The Washington Univ. Nematode EST Project, 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.wustl.edu
The library was constructed by Claire Murphy and Dr. James McCarter at Washington University, St. Louis. RNA was provided by Andrew Kloek of Divergence Inc., St. Louis, MO.
Putative full length read
The vector to vector length is 555
Seq primer: SLI primer

FEATURES

SOURCE

1..524
Location/Qualifiers
/organism="Pratylenchus penetrans"
/mol_type="mRNA"
/db_xref="taxon:45929"
/dev_stage="mixed stage"
/lab_host="DH10B (Invitrogen)"
/clone_lib="Pratylenchus penetrans mixed stage SLI TOPO v1"
/note="Vector: pCRII-TOPO (Invitrogen); Site_1: EcorI; Site_2: EcorI; The library was constructed by Claire Murphy and Dr. James McCarter at Washington University, St. Louis. Oligo(dT)-SLI PCR based library. cDNA PCR products of size >400 nucleotides containing SLI on the 5' end and oligo(dT) on the 3' end were non-directionally cloned into pCRII-TOPO(Invitrogen) following the Topo TA cloning protocol. RNA was provided by Andrew Kloek of Divergence, Inc., St. Louis, MO."

ORIGIN

Query Match 68.6%; Score 19.2; DB 3; Length 524;
Best Local Similarity 87.5%; Pred. No. 1.5e+03;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GAACAGCAATTCACAAACGCTTG 24
Db 102 GAGCAGCAATTCACAAACGCTTG 125

RESULT 46
BO626527 573 bp mRNA linear EST 01-JUL-2002
LOCUS p31d01.y1 Pratylenchus penetrans mixed stage SLI TOPO v1

DEFINITION Pratylenchus penetrans cDNA 5' similar to SW:RS15_HUMAN P11174 40S
RIBOSOMAL PROTEIN S15 ; , mRNA sequence.

BO626527
BO626527.1 GI:21653705

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

McCarter, J., Clifton, S., Chiapelli, B., Pape, D., Martin, J., Wylie, T., Dante, M., Marra, M., Hillier, L., Kucaba, T., Theising, B., Bowers, Y., Gibbons, M., Ritzer, E., Bennett, J., Franklin, C., Tsagarishevili, R., Ronko, I., Kennedy, S., Maguire, L., Beck, C., Underwood, K., Steptoe, M., Allen, M., Person, B., Swaller, T., Harvey, N., Schurk, R., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.
The Washington Univ. Nematode EST Project, 1999
Unpublished (1999)
Contact: McCarter JP
The Washington Univ. Nematode EST Project, 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.wustl.edu
The library was constructed by Claire Murphy and Dr. James McCarter at Washington University, St. Louis. RNA was provided by Andrew Kloek of Divergence Inc., St. Louis, MO.
Putative full length read
The vector to vector length is 574
Seq primer: SLI primer
High quality sequence stop: 525.

FEATURES

SOURCE

1..573
Location/Qualifiers
/organism="Pratylenchus penetrans"
/mol_type="mRNA"
/db_xref="taxon:45929"
/dev_stage="mixed stage"
/lab_host="DH10B (Invitrogen)"
/clone_lib="Pratylenchus penetrans mixed stage SLI TOPO v1"
/note="Vector: pCRII-TOPO (Invitrogen); Site_1: EcorI; Site_2: EcorI; The library was constructed by Claire Murphy and Dr. James McCarter at Washington University, St. Louis. Oligo(dT)-SLI PCR based library. cDNA PCR products of size >400 nucleotides containing SLI on the 5' end and oligo(dT) on the 3' end were non-directionally cloned into pCRII-TOPO(Invitrogen) following the Topo TA cloning protocol. RNA was provided by Andrew Kloek of Divergence, Inc., St. Louis, MO."

ORIGIN

Query Match 68.6%; Score 19.2; DB 3; Length 573;
Best Local Similarity 87.5%; Pred. No. 1.5e+03;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GAACAGCAATTCACAAACGCTTG 24

Db 121 GAGCAGCAATTCACAAACGTTG 144

RESULT 47
LOCUS BU220286
DEFINITION BU220286 707 bp mRNA linear EST 25-NOV-2002
603107182F1 CSEBCHN04 Gallus gallus cDNA clone CHEST48C21 5', mRNA
sequence.
ACCESSION BU220286
VERSION BU220286.1 GI:25405987
KEYWORDS EST.
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
1 (bases 1 to 707)
Boatman, P.B., Sanz-Ezquerro, J., Overton, I.M., Burt, D.W., Bosch, E.,
Fong, W.T., Tickle, C., Brown, W.R.A., Wilson, S.A. and Hubbard, S.J.,
A Comprehensive Collection of Chicken CDNA
Curr. Biol. 12 (22), 1965-1969 (2002)
12445392

REFERENCE
AUTHORS Contact: Simon Hubbard
Department of Biomolecular Sciences
University of Manchester Institute of Science and Technology
(UMIST)
PO Box 88, Manchester, M60 1QD, UK
Tel: 01612008930
Fax: 01612360409
Email: Simon.Hubbard@umist.ac.uk.
Location/Qualifiers

FEATURES
source
1..707
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="White Leghorn, Hisex"
/db_xref="taxon:9031"
/clone="CHEST48C21"
/cissue_type="whole embryo"
/dev_stage="20-21"
/lab_host="DH10B"
/clone_lib="CSEBCHN04"
/note="Organ: whole embryo; Vector: pBluescript II KS(+);
Site 1: EcoRI; Site 2: NotI; This normalized library was
constructed from 1 million independent clones. cDNA
synthesis was initiated using an oligo(dT) primer, using
methylated C in the first strand synthesis reaction.
Following this first strand reaction, double-stranded cDNA
was blunt-ended, ligated to NotI adapters, digested with
EcoRI, size-selected, and cloned into the NotI and EcoRI
compatible sites of a custom modified MCS of the
pBluescript (KS+) vector. The library was normalized in 2
rounds using conditions adapted from Soares et al., PNAS
(1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6
(1996): 791, except that a significantly longer
reannealing hybridization was used."

ORIGIN
Query Match 68.6%; Score 19.2; DB 3; Length 707;
Best Local Similarity 87.5%; Pred. No. 1.5e+03;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GAACAGCAATTCACAAACGTTG 24
|||||
Db 329 GAACAGCATATACAAACGTTG 352
|||||

RESULT 48
LOCUS DR154444/c
DEFINITION DR154444 783 bp mRNA linear EST 16-JUN-2005
49110365 Drosophila pseudoobscura embryonic cDNA library Drosophila
pseudoobscura cDNA clone K19 3', mRNA sequence.
ACCESSION DR154444

VERSION DR154444.1 GI:67899534
KEYWORDS EST.
SOURCE Drosophila pseudoobscura
ORGANISM Drosophila pseudoobscura
Bukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachyera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 783)
Richard, S., Liu, Y., Bettencourt, B.R., Hradecky, P., Letovsky, S.,
Nielsen, R., Thornton, K., Hubisz, M.J., Chen, R., Meisel, R.P.,
Couronne, O., Hua, S., Smith, M.A., Zhang, P., Liu, J., Bussemaker, H.J.,
van Batenburg, M.F., Howells, S.L., Scherer, S.B., Sodergren, E.,
Matthews, B.B., Crosby, M.A., Schroeder, A.J., Ortiz-Barrientos, D.,
Rives, C.M., Metzker, M.L., Muzny, D.M., Scott, G., Steffen, D.,
Weiler, D.A., Worley, K.C., Havlak, P., Dublin, K.J., Egan, A.,
Gill, R., Hume, J., Morgan, M.B., Miner, G., Hamilton, C., Huang, Y.,
Malden, L., Verdugo, D., Clerc, Blankenburg, K.P., Dubchak, I.,
Noor, M.A., Anderson, W., White, K.P., Clark, A.G., Schaeffer, S.W.,
Gelbart, W., Weinstock, G.M. and Gibbs, R.A.
Comparative genome sequencing of Drosophila pseudoobscura:
chromosomal, gene, and cis-element evolution
Genome Res. 15 (1), 1-18 (2005)
15632085

REFERENCE
AUTHORS Contact: Stephen Richards
Human Genome Sequencing Center
Baylor College of Medicine
One Baylor Plaza, Houston, TX 77030, USA
Tel: 713-798-6667
Email: stephen@bcm.tmc.edu
NCBI Trace Archive: 22669758
Insert Length: 1750 Std Error: 0.25.
Location/Qualifiers

FEATURES
source
1..783
/organism="Drosophila pseudoobscura"
/mol_type="mRNA"
/db_xref="taxon:7237"
/clone="K19"
/dev_stage="0-18h embryos"
/clone_lib="Drosophila pseudoobscura embryonic cDNA
library"
/note="Vector: pOTB7; Site 1: EcoRI; Site 2: XhoI; oligo
dt priming from poly A+ RNA, directionally cloned"

ORIGIN
Query Match 68.6%; Score 19.2; DB 9; Length 783;
Best Local Similarity 87.5%; Pred. No. 1.5e+03;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GAACAGCAATTCACAAACGTTG 24
|||||
Db 764 GAACACACACACAAACGTTG 741
|||||

RESULT 49
LOCUS BI153495
DEFINITION 602915583F1 NCI_CGAP_Lu29 Mus musculus cDNA clone IMAGE:5066108 5',
mRNA sequence.
ACCESSION BI153495
VERSION BI153495.1 GI:14613496
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 817)
NIH-MGC http://mgc.nci.nih.gov/.
NATIONAL INSTITUTES OF HEALTH, MAMMALIAN GENE COLLECTION (MGC)
UNPUBLISHED (1999)
CONTACT: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.

CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
<http://image.lnl.gov>
 plate: L1AM1178 row: e column: 21
 High quality sequence start: 29
 High quality sequence stop: 815.

FEATURES

source

1. .817
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="Czech II"
 /db_xref="taxon:10090"
 /clone="IMAGE:5066108"
 /tissue_type="spontaneous tumor, metastatic to mammary."
 /stem_cell_origin="Stem cell origin."
 /lab_host="DH10B"
 /clone_lib="NCI CGAP Lx29"
 /note="Organ: lung; Vector: pCMV-SPORT6; Site_1: Salt;
 Site_2: Not1; Cloned unidirectionally. Primer: Oligo dT.
 Library constructed by Life Technologies. Investigator
 providing samples: Gilbert Smith, NIH"

ORIGIN

Query Match 68.6%; Score 19.2; DB 2; Length 817;
 Best Local Similarity 87.5%; Pred. No. 1.5e+03;
 Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AACGACATACACAAACCGTTGTG 25

Db 664 AACGACATACACAAACCGTTGTG 687

RESULT 50

CT474369 818 bp DNA linear GSS 03-NOV-2005
 Sus scrofa genomic clone CH242-469H10, genomic survey sequence.

CT474369
 CT474369.1 GI:80214467

KEYWORDS

GSS.

SOURCE

Sus scrofa (pig)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Suidae;
 Sus.

1 (bases 1 to 818)
 Humphray,S.J., Plumb,R.W. and Durham,J.L.

Direct Submission
 Submitted (01-NOV-2005) The Sanger Institute, Wellcome Trust Genome
 Campus, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humphrey@sanger.ac.uk Unpublished

COMMENT
 This sequence was generated from the T7 end of BAC 469H10. 469H10
 is part of the CHOBI-242 BAC library created by P. de Jong. Further
 details: http://www.sanger.ac.uk/Projects/S_scrofa/.

FEATURES

source

1. .818
 /organism="Sus scrofa"
 /mol_type="genomic DNA"
 /db_xref="taxon:9823"
 /clone="CH242-469H10"
 /tissue_type="white blood cells"
 /note="vector pTARBAC1.3_BamHI
 sex female"

ORIGIN

Query Match 68.6%; Score 19.2; DB 14; Length 818;
 Best Local Similarity 87.5%; Pred. No. 1.5e+03;
 Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 5 AGCAATACAAACAACCGTTGTG 28

||||| ||||||| ||||| ||

Db 511 AGCAATACAAACAACCGTTGTG 534

Search completed: May 24, 2006, 07:54:55
 Job time: 1999.02 secs

THIS PAGE IS BLANK